Akron Law Review

Volume 57
Issue 2 *Intellectual Property Issue*

Article 1

2024

Governing Generative AI

Daryl Lim

Follow this and additional works at: https://ideaexchange.uakron.edu/akronlawreview

Part of the Intellectual Property Law Commons, and the Other Law Commons

Please take a moment to share how this work helps you through this survey. Your feedback will be important as we plan further development of our repository.

Recommended Citation
Lim, Daryl (2024) "Governing Generative AI," *Akron Law Review*: Vol. 57: Iss. 2, Article 1.
Available at: https://ideaexchange.uakron.edu/akronlawreview/vol57/iss2/1

This Article is brought to you for free and open access by Akron Law Journals at IdeaExchange@UAkron, the institutional repository of The University of Akron in Akron, Ohio, USA. It has been accepted for inclusion in Akron Law Review by an authorized administrator of IdeaExchange@UAkron. For more information, please contact mjon@uakron.edu, uapress@uakron.edu.

GOVERNING GENERATIVE AI

Daryl Lim[*]

I. Introduction 210
II. Governing Authorship 211
A. Anthropocentric Authorship 211
1. A Strict View of Originality 212
2. Originality with Chinese Characteristics 217
B. Creation's Evolution 221
1. Crisis in Hollywood 221
2. Tools and Co-Creators 226
III. Governing Infringement 228
A. Infringement by Ingestion 228
B. The Promise of Private Ordering and Its Limits 235
1. Blanket Licensing and Micro Licensing 235
2. Glaze and Nightshade 238
3. Indemnities and Public Relations 239
IV. Models of Governance 243
A. America's Market-Driven Approach 243
B. Europe's Right-Centered Approach 244

[*]H. Laddie Montague Jr. Chair in Law; Associate Dean for Research and Innovation; Founding Director, Intellectual Property Law and Innovation Initiative; Co-hire, Institute for Computational and Data Sciences; and Affiliate, Center for Socially Responsible Artificial Intelligence, Penn State University. This Article benefit from presentations at the 2023 UAIP Scholars Forum at the University of Akron School of Law, Pennsylvania Office of General Counsel University, Cumberland County Bar Association, Penn State Institute for Computational and Data Sciences Fall 2023 Symposium, Pennsylvania Bar Association Federal Practice Committee Annual Federal Practice Institute, UC Irvine Antitrust Law and Artificial Intelligence, Competition, Antitrust Law, and Innovation Forum, Antitrust - National Association of Attorneys Generals Eastern Regional Meeting, Fundamentals of Copyright in the Digital Era, Practicing Law Institute, the 7th Annual Texas A&M Intellectual Property Scholars Roundtable at Texas A&M University School of Law, the Texas A&M "A Decade Of Growth And Transformation In Intellectual Property And Technology Law" Symposium, and Georgetown Law Technology Review Symposium on "Artificial Lawyering — Law in the Age of Artificial Intelligence." I am grateful to Michael Doane, Elizabeth Townsend Gard, Camilla Hrdy, Eric Johnson, Emily Michiko Morris, Craig Nard, Mark Schultz, Betsy Rosenblatt, Peter Yu for their valuable comments and suggestions. This Article builds on my earlier piece, *AI & IP Innovation & Creativity in an Age of Accelerated Change*, 52 AKRON L. REV. 813 (2018).

Published by IdeaExchange@UAkron, 2024

C. Asian Pro-Innovation Approaches 247
V. Conclusion.....250

I. INTRODUCTION

In the burgeoning era of artificial intelligence (AI), the rise of generative AI technologies such as ChatGPT, Midjourney, and their ilk stands as a testament to the profound transformation sweeping across various creative domains. From the glitzy corridors of Hollywood to the bustling newsrooms of journalism, a palpable sense of crisis looms large, spurred by the advent of machines that can mimic and, at times, surpass human creativity. As these digital marvels churn out content with astonishing ease and flair, critical questions arise: What becomes of the human creator in this new algorithm-driven landscape? How should the law balance the right of creatives to compensation, consent, and control of their works against the profound technological progress generative AI offers? What approaches have leading economies around the world taken to resolve these issues?

Part I explores the impact of generative AI on creative industries, highlighting the crisis of confidence among creators due to the potential loss of control over their work. It discusses the traditional copyright framework that reserves authorship for humans, thus excluding AI-generated content from protection due to the lack of human creative involvement. This section also contrasts the conservative stance of the US copyright regime with more liberal interpretations in other jurisdictions, like China, where copyright protection for AI-generated works is recognized under certain conditions, emphasizing the need for adaptable legal frameworks to balance innovation with the protection of human creators' rights.

Part II addresses the complexities of copyright infringement in the context of generative AI, focusing on issues such as the unauthorized use of copyrighted materials for AI training and the exploration of private ordering mechanisms like licensing and indemnities as potential solutions. It highlights legal challenges posed by AI's ability to create derivative works without clear attribution to original human creators, thereby blurring the lines of copyright infringement. The section also delves into the effectiveness and limitations of private solutions such as Adobe's Firefly model, which seeks to navigate copyright compliance within AI-generated content and discusses innovative approaches like Glaze and

Nightshade that aim to protect creators' rights by manipulating data to hinder AI's unauthorized use. Additionally, it considers the role of indemnities in mitigating infringement risks for both AI developers and users, acknowledging the ongoing legal and ethical debates surrounding the fair use of copyrighted materials in the training and operation of AI systems.

Part III delves into the varying global approaches to AI regulation, highlighting the fragmented landscape that arises from different national policies and priorities. The section contrasts the United States' market-driven approach, which emphasizes innovation and free speech, with Europe's rights-focused strategy that provides mechanisms for copyright holders to control the use of their work for AI training. Additionally, it discusses Asia's pro-innovation stance, notably China's support for the copyrightability of AI-generated works amidst stringent regulations. The narrative underscores the challenges of creating a cohesive regulatory framework for a technology that inherently transcends borders, suggesting the need for international cooperation and the establishment of global norms. This part advocates for adaptable and forward-looking legal frameworks that balance innovation with the protection of human creators' intellectual property rights, emphasizing the importance of collaborative policymaking that involves a wide range of stakeholders.

II. GOVERNING AUTHORSHIP

Creators in various industries face an existential crisis due to the rise of generative AI technologies like ChatGPT, which challenge traditional notions of creativity and authorship. The section examines the legal conundrums posed by AI's creative capabilities, where copyright law traditionally reserves authorship to humans, thus excluding AI-generated content from protection. It contrasts the conservative approach of the US, which emphasizes human authorship, with more flexible interpretations in other jurisdictions such as China, suggesting a future where AI could be recognized as a co-creator.

A. Anthropocentric Authorship

There is mounting hype surrounding the disruptive potential of generative AI and the urgent need for governance frameworks.[1] From

1. *See How AI Image-Generators Work*, THE ECONOMIST (Jul. 10, 2023), https://www.economist.com/the-economist-explains/2023/07/10/how-ai-image-generators-work [https://perma.cc/W2NG-DVQH]. Explaining that generative AI is a type of deep learning, a software technique that uses layers of interconnected nodes that loosely mimic the structure of the human brain.

Hollywood to The New York Times, generative AI has sparked a crisis of confidence in creators who see technology as their hold on compensation, control, and consent over their content. The influx of works leveraging generative AI through newly accessible mainstream tools like ChatGPT, Midjourney, and Gemini has resurfaced the question of the degree of human creativity required for copyright protection. The policy debate here revolves around incentives. One view argues that incentives for generating AI works are unnecessary as altruistic and self-aggrandizing creators do require copyright protection.[2] However, the market-driven reality of content production, foundational to exclusive copyright rights, challenges the notion that incentives are not critical to human creation.

1. A Strict View of Originality

It is now clear that copyright law extends authorship only to humans, and AI-created works without human creative involvement or control cannot be copyrighted.[3] They are unprotectable not because it was AI and not the humans who supplied originality, but because humans did not make an original contribution to the work.[4] The Copyright Office requires registrants to exclude AI-generated content that is more than de minimis from registration applications as the purported human author neither created the relevant elements of the work independently nor supplied the necessary original, creative expression.[5] This strict interpretation of the

AI algorithms generate or create an output, such as text, a photo, a video, code, or data, from user instructions or prompts based on the datasets the AI tool has been trained on. A model that has learned which types of pixel arrangement correlate to the word "hippopotamus" can sample from its latent space to create a realistic image of the mammal. Adding more detail to the prompt—for example, "a renaissance-era oil painting of a green hippopotamus, somewhere along the river Nile"—requires the model to source additional layers of visual detail, such as image style, texture, color and location, and to combine them correctly.) Examples of generative AI tools include chatbots such as Google's Bard and OpenAI's ChatGPT; image generators such as Dall-E2, Midjourney and Stability AI. *See* Aditya Sharmam, TURING, *11 Best Generative AI Tools and Platforms*, https://www.turing.com/resources/generative-ai-tools; Harry Guinness, *The Best AI Image Generators In 2024*, ZAPIER (Feb. 22, 2024), https://zapier.com/blog/best-ai-image-generator/.

2. Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *4 (D.D.C. Aug. 18, 2023) ("Non-human actors need no incentivization with the promise of exclusive rights under United States law, and copyright was therefore not designed to reach them.").

3. *See id*. (finding that the Copyright Office did not act arbitrarily or capriciously in denying application for registration of autonomously generated work).

4. *Id*. at *3 ("Copyright is designed to adapt with the times. Underlying that adaptability, however, has been a consistent understanding that human creativity is the sine qua non at the core of copyrightability, even as that human creativity is channeled through new tools or into new media.").

5. A Rule by the Copyright Office, Library of Congress, *Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence*, (Mar. 16, 2023) https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-

law would be unremarkable on its own. However, in a global economic race for investment dollars, it is a self-inflicted handicap.

Congress designed copyright law to adapt and cover works created with evolving technologies.[6] It baked that malleability into the Copyright Act, which provides that copyright attaches to "original works of authorship fixed in any tangible medium of expression, now known or later developed."[7] A work "must be original, that is, the author's tangible expression of his ideas."[8] A work is original if it is independently created and possesses minimal creativity.[9] AI-generated images "absent any guiding human hand" are uncopyrightable.[10] Effort and economic value are irrelevant. Rather, intellectual labor, which indicates an indispensable human element of creation, is needed.[11]

In 1884, the Supreme Court ruled on the copyright of photographs, determining that despite being produced by a mechanical process, the photograph embodies the photographer's artistic decisions and thus warrants copyright protection.[12] One interpretation that follows is that the copyrightability of photographs rests on human creators' conceiving and designing the image and then using the camera to capture the image.[13] Where a work was created autonomously by a machine, and a registrant's claim to the copyright was only based on the fact of his ownership of the

guidance-works-containing-material-generated-by-artificial-intelligence [https://perma.cc/X9YN-H23F].

6. *See, e.g.*, Goldstein v. California, 412 U.S. 546 (1973) (explaining that the constitutional scope of Congress's power to "protect the 'Writings' of 'Authors'" is "broad," such that "writings" is not "limited to script or printed material," but rather encompasses "any physical rendering of the fruits of creative intellectual or aesthetic labor").

7. 17 U.S.C. § 102(a).

8. Mazer v. Stein, 347 U.S. 201 (1954); *see also id*. at 561 (defining "author" as "an 'originator,' 'he to whom anything owes its origin'").

9. Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991).

10. Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *4 (D.D.C. Aug. 18, 2023); *see also id.* ("The 'fixing' of the work in the tangible medium must be done 'by or under the authority of the author.'" [ ] In order to be eligible for copyright, then, a work must have an "author.").

11. *See, e.g.*, Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991) ("In Burrow-Giles, the Court . . . defined 'author,' in a constitutional sense, to mean 'he to whom anything owes its origin; originator; maker.' . . . As in The Trade-Mark Cases, the Court emphasized the creative component of originality. It described copyright as being limited to 'original intellectual conceptions of the author,' . . . and stressed the importance of requiring an author who accuses another of infringement to prove 'the existence of those facts of originality, of intellectual production, of thought, and conception.'").

12. Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 56 (1884).

13. *Id*. at 59–60 (The photographer develops a "mental conception" of the image together with decisions which may include "posing the [subject] in front of the camera, selecting and arranging the costume, draperies, and other various accessories in said photograph, arranging the subject so as to present graceful outlines, arranging and disposing the light and shade, suggesting and evoking the desired expression, and from such disposition, arrangement, or representation.").

machine, the registrant played no role in using the AI to generate the work.[14]

As one court recently put it, "The increased attenuation of human creativity from the actual generation of the final work will prompt challenging questions regarding how much human input is necessary to qualify the user of an AI system as an "author" of a generated work, the scope of the protection obtained over the resultant image, how to assess the originality of AI-generated works where the systems may have been trained on unknown preexisting works, how copyright might best be used to incentivize creative works involving AI, and more."[15] It concluded that "In the absence of any human involvement in the creation of the work, the clear and straightforward answer is the one given by the Register: No."[16] To put it differently, copyright law will not recognize authorship without a human in the creative process,

Courts previously declined to extend copyright protection to celestial beings, monkeys,[17] and gardens[18] because human authorship was lacking. AI-generated works, which are created autonomously by AI systems, are not covered by copyright protection. This is similar to the way ideas, facts, or content that is dictated by technical requirements rather than human creative effort are also not eligible for copyright.[19] Well, and good, but how much more human intervention is required?

According to the courts, not much. Anything with some creative spark could meet the bar for originality, no matter how crude, humble, or obvious.[20] The more advanced the technology and the more intelligent the tools are, the less human input is needed. Before cameras, people painted to reproduce images. Smartphone photos reflect the photographer's originality and intellectual investment and are protected by copyright law.

14. *See* Ex. E, First Request for Reconsideration at 2, ECF No. 13-5, at 2 ("It is correct that the present submission lacks traditional human authorship—it was autonomously generated by an AI."); Ex. G, Second Request for Reconsideration at 2, ECF No. 13-7, at 2 (same).

15. *Id*.

16. Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *6 (D.D.C. Aug. 18, 2023).

17. Naruto v. Slater, 888 F.3d 418 (9th Cir. 2018).

18. Kelley v. Chicago Park Dist., 635 F.3d 290, 304 (7th Cir. 2011) ("A 'living garden' was not copyrightable, in part because 'works owing their form to the forces of nature cannot be copyrighted.'" and "Creative elements of the garden "originate[d] in nature, not in the mind of the [human] gardener.").

19. *Thaler*, 2023 WL 5333236, at *6 (D.D.C. Aug. 18, 2023).

20. Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 345 (1991) ("[T]he requisite level of creativity is extremely low; even a slight amount will suffice. The vast majority of works make the grade quite easily, as they possess some creative spark, 'no matter how crude, humble or obvious' it might be.").

Generative AI works similarly in response to a prompt from a human, determined and executed by the technology, not the human user, functioning like instructions to a commissioned artist, even when prompts are iterative.

As a policy matter, copyright law is utilitarian.[21] The creator is a means to the goal of producing and disseminating creative content,[22] and the law meters the protection it gives to them.[23] The law prescribes the threshold for granting a copyright, scope, and duration.[24] It also protects the public's interests through statutory and common law defenses. For example, fair use allows the use of copyrighted material for research, current affairs reporting, and criticism.[25] Fair use also allows rivals to use copyrighted content, even when averse to the copyright owner's commercial interests.[26] Competition between firms vying to succeed in the marketplace can lower prices, improve the quality of goods or services, increase firms' productivity, spur the introduction of new products, and otherwise motivate innovation.[27]

Business needs to drive protection, as seen in the work-for-hire doctrine, a legal fiction that allows non-human ownership of human-authored works.[28] Denying copyright impacts creators' ability to protect

21. James A.D. White, *Misuse or Fair Use: That Is the Software Copyright Question*, 12 BERKELEY TECH. L. J. 251, 254 (1997) (discussing the utilitarian goals of IP policy).

22. Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 349 (1991) ("The primary objective of copyright is not to reward the labor of authors, but '[t]o promote the Progress of Science and useful Arts.' Art. I, § 8, cl. 8. To this end, copyright assures authors the right to their original expression, but encourages others to build freely upon the ideas and information conveyed by a work.").

23. *See id.* at 350 ("As applied to a factual compilation, assuming the absence of original written expression, only the compiler's selection and arrangement may be protected; the raw facts may be copied at will. This result is neither unfair nor unfortunate. It is the means by which copyright advances the progress of science and art.").

24. *See* Chapter 1 of Title 17 of the U.S. Code. Outlining the subject matter and scope of copyright, including definitions, the subject matter of copyright, compilations and derivative works, and limitations on exclusive rights. *See also* 17 U.S.C. § 302 (providing the standard duration of the author's life plus 70 years, and specific terms for joint works, anonymous works, pseudonymous works, and works made for hire.).

25. 17 U.S.C. § 107.

26. Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183, 1204 (2021) (observing that "many common fair uses are indisputably commercial").

27. U.S. DEP'T OF JUSTICE & FED. TRADE COMM'N, ANTITRUST ENFORCEMENT AND INTELLECTUAL PROPERTY RIGHTS: PROMOTING INNOVATION AND COMPETITION 1 (2007), https://www.justice.gov/atr/antitrust-enforcement-and-intellectual-property-rights-promoting-innovation-and-competition [https://perma.cc/8KTC-KJUG].

28. Thaler v. Perlmutter, No. CV 22-1564 (BAH), 2023 WL 5333236, at *6 (D.D.C. Aug. 18, 2023) ("The work-for-hire provisions of the Copyright Act, too, presuppose that an interest exists to be claimed."); *see* 17 U.S.C. § 201(b) ("In the case of a work made for hire, the employer . . . owns all of the rights comprised in the copyright.").

their creative works. They cannot sue for infringement or recover statutory damages and attorney's fees under the Copyright Act.[29] As I noted in an earlier work, "Excluding AI-generated work for protection leaves an economic lacuna, which US IP policy is loath to do. Allowing AI-generated works to fall into the public domain reduces the incentive to invest in the growth of the industry."[30] In this light, a strict interpretation of authorship is out of step with copyright's goals.

The Copyright Office's refusal to recognize copyright in "Théâtre D'opéra Spatial" is a case in point.[31] Jason Allen used 624 text prompts to create Théâtre D'opéra Spatial with Midjourney. The Copyright Board found that the artwork contained "more than a de minimis amount of content generated by artificial intelligence," which must be disclaimed to satisfy the human authorship requirement and qualify for copyright registration.[32] According to the Board, inputting commands does not amount to human authorship because the "traditional elements of authorship" are determined and executed by AI, not the human user. The selection and arrangement of prompts and parameters failed to meet the Copyright Office's standard of "artfully arranged" content or modifications.[33]

The Board noted that an AI system does not carry out instructions to create a specific expressive result because it does not understand

29. Adrian F. Vallens, Nicolas A. Jampol & Farrah Vazquez, *Appreciable or de Minimis: That Is the AI Question*, LEXOLOGY (Dec. 4, 2023) https://www.lexology.com/library/detail.aspx?g=4dcdc0bc-0dd8-43f3-ba09-1e5c9f7c26a3 [https://perma.cc/34V8-DNH5].

30. Daryl Lim, *AI & IP: Innovation & Creativity in an Age of Accelerated Change*, 52 AKRON L. REV. 813, 841 (2018); *see also id*. at 843 ("With AI-generated works, AI is simply a means by which one human creates in a way specified by another. In the analog world, Spirographs, recipes, and blueprints allow us to create according to another's specified instructions.").

31. Copyright Review Board, *Second Request for Reconsideration for Refusal to Register Théâtre D'opéra Spatial*, (Sep. 5, 2023) (SR # 1-11743923581; Correspondence ID: 1-5T5320R) https://www.copyright.gov/rulings-filings/review-board/docs/Theatre-Dopera-Spatial.pdf [https://perma.cc/WNZ7-T8PH].

32. *See Copyright Registration Guidance*: *Works Containing Material Generated by Artificial Intelligence*, 88 Fed. Reg. 51 at 16191 (Mar. 16, 2023), https://www.federalregister.gov/documents/2023/03/16/2023-05321/copyright-registration-guidance-works-containing-material-generated-by-artificial-intelligence [https://perma.cc/NNL4-4WB6].

33. *Id*. To determine whether disclosure is necessary, the Copyright Office instructs applicants to answer the following questions: (1) Would the AI-generated material, standing on its own, be copyrightable if it had been created by a human author?, (2) If so, then a brief statement disclosing the material must be included in the copyright application, and the AI-generated material will be excluded from registration. *Id.* However, if the work does not contain enough creativity to support a copyright, there is no need to disclose it, and it is not disclaimed. *Id.*

instructions the way humans do.[34] There could be thousands of ways it could translate the prompt into an image, almost in a way that is "random," so one cannot predict what the image it generates will look like. It needed hundreds of iterations of prompts and images before the prompter finally found four that were satisfactory and captured what he wanted.[35] In other words, the style and other elements of authorship were determined by the AI system and not the user.

2. Originality with Chinese Characteristics

In contrast to the US, China has embraced the more liberal interpretation of attributing AI-generated work to prompters.[36] The Beijing Internet Court had to consider if an image of a young Asian lady generated using 300 prompts qualified as a copyrighted work.[37] Strikingly, the user argued that the "intellectual investment" of selecting and arranging the inputs necessary to generate the image at issue is akin to photographers manually adjusting antique cameras to achieve a desired result in a photograph.[38]

The court agreed, finding the work original, expressive, and an "intellectual achievement."[39] The "originality" flowed from the

34. Copyright Review Board, *supra* note 31, at 16192.

35. *Id*. ("[W]hen an AI technology receives solely a prompt from a human and produces complex written, visual, or musical works in response, the 'traditional elements of authorship' are determined and executed by the technology—not the human user.").

36. *See* Fredrick Xie, *For the First Time AI Generated Photo Gets Copyright In China*, MONDAQ (Dec. 7, 2023) (characterizing the US position on the issue as "a contrary decision" that works generated entirely by AI are uncopyrightable for lack of human authorship); J.D. Capelouto, *How China's $70 AI Copyright Ruling Impacts The World*, SEMAFOR (Dec 4, 2023), https://www.semafor.com/article/12/04/2023/china-ai-copyright-court-ruling [https://perma.cc/339H-Z4BL] (noting that the ruling "differs from a decision by the U.S. Copyright Office" and noting that "[t]he finding, the first of its kind in China, is in direct conflict with the human authorship requirement under U.S. copyright law").

37. Tamara Céline Winegust, *Acquiring and Licensing Content in the Age of GenerativeAI*, LEXOLOGY (Dec. 19, 2023), https://www.lexology.com/library/detail.aspx?g=a1ab3481-3915-4164-aa54-8d770f7b98c8 [https://perma.cc/98DM-RPA8].

38. Tian Lu, *Chinese Court Deems AI-Generated Image Has Copyright – Assessing The Possibly Over-Hasty 'Spring Breeze' Case*, THE IPKAT (Dec. 27, 2023) https://ipkitten.blogspot.com/2023/12/chinese-court-deems-ai-generated-image.html [https://perma.cc/5AYU-GN4J].

39. *Id*. ("The image generated is not a ready-made image called up by a search engine, nor is it a combination of various elements predefined by the software designer. In layman's terms, the role or function of the Stable Diffusion model is similar to that of human beings who have acquired some abilities and skills through learning and accumulation."); *see also id*. ("In the entire process, from the plaintiff's conception to the final selection of the image in question, the plaintiff has made a certain intellectual input, such as designing the presentation of the characters, choosing the prompts, arranging the order of the prompts, setting the relevant parameters, and selecting the image that meets

continuous prompts and repeated parameter adjustments to create a picture that reflected the user's "aesthetic choice and personalized judgment."[40] According to the court, the image had not existed previously, nor was it merely a combination of preexisting elements.[41] Instead, similar to a human who develops skills and knowledge over time, the algorithm creates images in response to text descriptions provided by users, effectively replacing the human act of sketching and coloring, thereby materially manifesting human creativity and ideas.[42]

Significantly, the court accepted the AI user's analogy of generative AI to a photographer as the author of the resulting copyright-protected photograph, designating them as tools for humans, avoiding the tangle of parsing copyrightable from noncopyrightable content.[43] Like a paintbrush, AI is a tool for human creation by assuming the human's role of drawing lines, filling in colors, and presenting the human's ideas and creativity in a tangible medium.[44]

---

his expectations. The image in question reflected the plaintiff's intellectual input, so the requirement of "intellectual achievement" has been satisfied.").

40. *Id*. (holding that "originality' requires that a work is created independently by the author and reflects the author's individual expression. 'Mechanical intellectual achievements' should be excluded. For example, works completed following a specific sequence, formula, or structure will lead to the same results for different people. Such works, due to their uniqueness in expression, lack originality.").

41. *Id*. ("It can be said that the lines and colors that constitute the image in question were basically 'drawn' by the Stable Diffusion model, which is different from what people have used before.").

42. *Id*. ("However, the plaintiff designed the figure and its presentation, as well as other elements of the image, through the prompts. He also set the parameters for the layout and composition of the image, reflecting the plaintiff's choice and arrangement. On the other hand, the plaintiff, through the input prompts, set the relevant parameters, obtained the first image, continued to add prompts, modified the parameters, continued to adjust and amend, and finally obtained the image. This adjustment and amendment process also reflects the plaintiff's aesthetic choice and personalized judgment.").

43. *Id.* ("Nowadays, the camera functions of smartphones are becoming increasingly powerful and easier to use. However, as long as smartphone photos reflect the photographer's creative skills, they still constitute a photographic work and are protected by copyright law. As technology develops and the relevant tools become smarter, the human input decreases. This, nevertheless, does not affect our continued use of the copyright system to encourage the creation of works."); *but see* commentary by Tian Lu, at *id*. ("Stable Diffusion operates through a mechanism vastly different from cameras. Imagine a large black box filled with numerous images, each coded with a specific set of information, including values and random seeds. In an abstract sense, AA's process of 'creation' was akin to reaching into a general area of this black box and getting a pre-prepared image located in that area. Regardless of what exactly the hand gesture is, it does not 'hold a pen' and thus does not 'draw' an image in the black box.").

44. *Id*. ("At this stage, generative AI models do not have free will and are not subjects in law. In essence, people still use tools to create, that is, humans make intellectual input in the entire creation process, not AI models.").

Evoking "work for hire" reasoning, the court analogized AI to a commissioned artist, noting that the artist has free will in selection and judgment and, therefore, a creator of the work; AI does not and is, therefore, a tool for the human using it to create a copyrightable work.[45] Indeed, the work hire doctrine supports the proposition that an individual's "creativity" can be extremely insignificant and still be deemed the author. To underscore the utilitarian copyright policy guiding the court's judgment, it wrote:

> Currently, a new generation of generative AI technology is increasingly being used. The Stable Diffusion model and models with similar functions can claim beautiful images based on text descriptions. Many people, including those without drawing skills, are trying to use these new models to generate content and present their ideas and designs in a tangible way, greatly improving the efficiency of creating images. It should be said that generative AI technology has changed the way people create; that is the same impact that many technological advances have had throughout history. The development process of technology is the process of gradually outsourcing human work to machines.
>
> Encouraging creativity is widely recognized as the core purpose of the copyright system. Only by correctly applying the copyright system, using appropriate legal means, and encouraging more people to use the latest tools to create can it be more conducive to the creation of works and the development of AI technology. Under this background and technological reality, the images generated by AI, if they can reflect human's original intellectual input, should be recognized as works and enjoy protection under copyright law.[46]

Ironically, a court in a communist country recognized the market-driven reality of content production, foundational to exclusive copyright rights, and challenged the notion that incentives are not critical to human creation. The strictness of the US approach on the issue of copyrightability of AI-generated works has not passed unobserved by stakeholders. As Laura Lamansky noted:

> The US has taken the strictest stance on artificial intelligence laws to date compared to the international community. US courts and the US Copyright Office have determined that only human-authored creative works are eligible for copyright protection. Any works created by AI would not qualify for such protections. The US Copyright Office has even gone so far as to require copyright applicants to disclose what

45. *Id.*
46. *Id.*

> portions of the work were created by AI. Meanwhile, China's courts have swung the other way - allowing copyright protection for AI-generated images. The US's stricter approach - made even stricter by the requirement of registration prior to suing to enforce a copyright right - may lead to stagnation of AI-generated works and, as a result, be less welcoming to the businesses that develop AI technologies.[47]

The Chinese approach has much to commend it. Recognizing the authorship of prompters creates a legally cognizable path to takedown mechanisms. After all, a camera is a machine. It isn't easy to see why a smartphone photo of fireworks is copyrightable, but prompting Midjourney to produce a photograph of fireworks might not be. The alternative would have paved the way for others to appropriate generative AI works without accountability. Without a legal structure that incentivizes the creation of AI, society would miss out on these potential benefits. China's exploding AI market is projected to exceed $26 billion by 2026.[48] Chinese generative AI will contribute $4.2 trillion worth of economic value out of the industry's 90 trillion yuan globally by 2035.[49] It is a slice of the AI pie Americans risk losing out on.

In contrast to the conclusion reached by the Copyright Office, the US Patent and Trademark Office recognizes that humans using AI as a tool in making an invention may still qualify for inventorship.[50] An invention created by AI can be patented as long as a human meets the criteria for being a co-inventor by making a significant contribution to the invention.[51] Even in cases where AI contributes more significantly, a human can be recognized as the sole inventor and initial owner if they have made a substantial contribution. AI models are predominantly under

47. Laura Lamansky, *AI Patchwork of Laws: How the International Community is Shaping AI Laws & Regulations* (Jan. 10, 2024), https://insights.michaelbest.com/post/102iwpa/ai-patchwork-of-laws-how-the-international-community-is-shaping-ai-laws-regula [https://perma.cc/EF5U-FYYR].

48. Intellectual property library 2023-11-28 11:02 Posted on Zhejiang https://mp.weixin.qq.com/s?__biz=MzAwNDE3MjA5NA==&mid=2677385275&idx=1&sn=a8ccdbb118604473d8fd198f82df7e30 [https://perma.cc/2MVP-5XLN].

49. Ben Jiang, *Chinese Generative AI Will Contribute $4.2 Trillion Worth Of Economic Value Out Of The Industry's 90 Trillion Yuan Globally By 2035*, SOUTH CHINA MORNING POST (Dec. 15, 2023) https://www.scmp.com/tech/tech-trends/article/3245217/chinese-generative-ai-account-third-industrys-economic-value-2035-beijing-think-tank-says [https://perma.cc/FS6X-WVJW].

50. U.S. Patent & Trademark Office, *Inventorship Guidance for AI-Assisted Inventions,* (Feb. 13, 2024) https://www.federalregister.gov/documents/2024/02/13/2024-02623/inventorship-guidance-for-ai-assisted-inventions [https://perma.cc/XV3D-PGVY].

51. *Id.*

tight human supervision and direction rather than emerging autonomously from the AI's initiatives or vague instructions.[52]

Generative AI has the potential to transform the knowledge economy. Failing to reward creatives with authorship could hinder the deployment of new AI tools and render millions of jobs unsustainable. On the other hand, giving creatives leeway to bludgeon developers with lawsuits in the name of protecting their economic expectations risks stymying technological innovation. The key to navigating the rules of copyright enforcement in this space requires an understanding of how the enterprise of creation itself will evolve with generative AI.

B. Creation's Evolution

There are seismic shifts in creative processes due to generative AI. The section portrays Hollywood's existential angst as AI threatens traditional roles, with performers fearing replacement by synthetic entities, illustrating a broader creative upheaval. It further explores the dual role of AI as both a tool and a collaborator, democratizing creativity and challenging conventional authorship paradigms, thereby redefining the landscape of artistic creation and intellectual property.

1. Crisis in Hollywood

In "The Great Automatic Grammatizator," Roald Dahl wrote about a machine capable of writing short stories and novels at a button, displacing human writers in its ceaseless mechanical production.[53] Today, the future of work seems imperiled. A recent Gallup poll found that 75% of US adults believe AI will lead to fewer jobs.[54] The threat of AI coming for jobs in Hollywood put it at a standstill for half a year.[55] Performers feared they could soon be replaced by entirely synthetic "actors" generated by AI using the body parts of many different humans, whose

52. Dennis Crouch, *Joint Inventorship: AI-Human Style* (Feb. 12, 2024), https://cdn.patentlyo.com/media/2024/02/Inventorship-Guidance-for-AI-Assisted-Inventions.pdf [https://perma.cc/C6AC-UJDW] ("Opponents could contend that full conception remains legally required for inventorship and that this expansion of the inventorship doctrine exceeds the statutory language. It is not clear who will have standing to make this particular argument.").

53. ROALD DAHL, THE GREAT AUTOMATIC GRAMMATIZATOR AND OTHER STORIES (2001).

54. Stephanie Marken & Tara Nicola, *Three in Four Americans Believe AI Will Reduce Jobs* (Sep. 13, 2023), https://news.gallup.com/opinion/gallup/510635/three-four-americans-believe-reduce-jobs.aspx,%20accessed%2001/27/2023 [https://perma.cc/87XM-VQF5].

55. Danielle Broadway & Lisa Richwine, *Actors Ratify Three-Year Contract, Ending Hollywood's Labor Turmoil*, REUTERS (Dec. 6, 2023), https://www.reuters.com/business/media-telecom/hollywood-actors-union-ratifies-three-year-contract-after-strike-2023-12-06/ [https://perma.cc/QN99-WSND].

likenesses have been scraped from film archives.[56] Writers fear movie studios using AI to churn out blockbuster scripts and publishers using them to churn out books.[57] The contagion of fear affecting Hollywood is symptomatic of a broader epidemic gale of creative disruption sweeping across every field of endeavor.

History offers two key insights. Firstly, technological advancements have consistently reshaped established norms. The invention of the printing press catalyzed the creation of the first copyright laws in the 18th century.[58] The advent of television in the 1960s led to adjusted royalty payments for film actors whose works were broadcast on TV.[59] Search engines, too, faced lawsuits upon their introduction.[60]

Similarly, the digital music revolution initiated by platforms like Napster eventually resulted in formal agreements between streaming services and record labels.[61] This shift was driven by consumers' demand for a legitimate and convenient way to access music, leading to the rise of streaming platforms like Spotify, which offered a legal alternative to the cumbersome process of searching for music online, transitioning the music industry from physical formats like CDs to digital ones.[62] The focus was not on evading payment but on the ease of accessing beloved

56. Matt O'Brien, *Artists Push for AI Copyright Reforms But Tech Industry Isn't In A Rush*, SEATTLE TIMES (Nov. 27, 2023), https://www.seattletimes.com/business/please-regulate-ai-artists-push-for-u-s-copyright-reforms-but-tech-industry-says-not-so-fast/ [https://perma.cc/D4Z8-4PH2] ("Country singers, romance novelists, video game artists and voice actors are appealing to the U.S. government for relief — as soon as possible — from the threat that artificial intelligence poses to their livelihoods.").

57. The Authors Guild recently reported that 70 percent of authors believe publishers will replace human writers with AI-generated books.

58. *Copyright Timeline: A History of Copyright in the United States*, ASS'N OF RES. LIBRARIES, https://www.arl.org/copyright-timeline/ [https://perma.cc/Y7GY-RVF5] ("in 1710 Parliament enacted the Statute of Anne to address the concerns of English booksellers and printers.").

59. Andrew Dalton, *What residuals are — and why Hollywood actors and writers are striking over them*, KCRA (July 20, 2023), https://www.kcra.com/article/what-are-residuals/44601388 [https://perma.cc/G9A4-P37S].

60. *See, e.g.,* Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003); Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007).

61. *See generally*, A&M Recs., Inc. v. Napster, Inc., 284 F.3d 1091 (9th Cir. 2002).

62. Louis Menand, *Is A.I. the Death of IP?*, THE NEW YORKER (Jan. 15, 2024), https://www.newyorker.com/magazine/2024/01/22/who-owns-this-sentence-a-history-of-copyrights-and-wrongs-david-bellos-alexandre-montagu-book-review [https://perma.cc/VL7D-LQS5] ("Napster, launched in 1999, is the landmark example. Its peer-to-peer file-sharing system was determined to be piracy, but Napster still revolutionized the music industry by moving it into the streaming business. Performance revenue aside, music income now comes primarily not from CD sales but from licensing deals. Spotify is a direct descendant of the Napster case.").

music.[63] Despite initial resistance, royalty systems have been established and functioned effectively[64]

The second insight is that dramatic reactions and hyperbolic predictions have been common responses to new technologies.[65] French artist Paul Delaroche famously proclaimed painting obsolete upon encountering the daguerreotype, an early form of photography, yet painting adapted and thrived, liberated from the need for strict realism.[66] It took until the latter half of the 19th century for photography to be recognized as copyrightable.[67]

Early cinema and radio were feared to undermine live performance, yet they paved the way for the emergence of superstars, a phenomenon that became prominent in the 1920s.[68] Television's arrival sparked fears and strikes in Hollywood in 1960, yet the film industry has continued to grow, with blockbuster films claiming a larger share of box office revenue.[69] In the internet era, despite the vast array of user-generated content on platforms like YouTube and TikTok, top hits by major artists have gained even greater prominence, illustrating the ongoing evolution and adaptation of creative industries in the face of technological change.[70]

Content creators who adapt to new monetization strategies are set to flourish. The count of musicians making over $1,000 annually from Spotify royalties has doubled in the last six years, while those earning

63. *See* Caroline Day, *AI Observer, Issue 4: What if Generative AI does infringe copyright? A lesson from history*, LEXOLOGY, (Nov. 9, 2023), https://www.hlk-ip.com/ai-observer-issue-4-what-if-generative-ai-does-infringe-copyright/ [https://perma.cc/KGR9-3AUR].

64. *See* Ben Sisario, *Musicians Say Streaming Doesn't Pay. Can the Industry Change*?, N.Y. TIMES (May 7, 2021), https://www.nytimes.com/2021/05/07/arts/music/streaming-music-payments.html.

65. *Cf., e.g.*, John Maynard Keynes, *Economic Possibilities for Our Grandchildren*, reprinted in ESSAYS IN PERSUASION 358, 369 (1930) (explaining that, in the near future, the need for steady work would be eliminated, "For three hours a day is quite enough to satisfy the old Adam in most of us!").

66. Caterina Bellinetti, *"From Today Painting is Dead": Photography's Revolutionary Effect*, ART & OBJECT (Apr. 8, 2019), https://www.artandobject.com/news/today-painting-dead-photographys-revolutionary-effect [https://perma.cc/9DUW-CMCC].

67. Menand, *supra* note 62.

68. *The Dawn of the Omnistar*, THE ECONOMIST (Nov. 9, 2023), https://www.economist.com/leaders/2023/11/09/how-artificial-intelligence-will-transform-fame [https://perma.cc/H7UA-27RP].

69. Thomas Doherty, *The Last Time Actors and Writers Both Went on Strike: How Hollywood Ended the 1960 Crisis*, THE HOLLYWOOD REP. (July 18, 2023, 6:45 AM) https://www.hollywoodreporter.com/business/business-news/sag-wga-1960-hollywood-strike-reagan-history-1235538551/ [https://perma.cc/LE5S-GH2P].

70. *The Dawn of the Omnistar*, *supra* note 68.

above $10 million annually have increased fivefold.[71] AI-driven dubbing technologies now enable actors and podcasters to reach international audiences in their voices instantly.[72] Soon, video editing to sync lip movements with dubbed languages will become the norm. Digital enhancements will extend the careers of actors and even allow for performances after their passing.[73] Additionally, celebrities will have opportunities to engage with their audience through emerging new formats.[74]

Value differentiation between content providers that deploy generative AI and those that do not will grow, given rapid content creation capabilities and new market entrants.[75] Changes from technological revolutions embodied Schumpeter's gale of 'creative destruction' that swept away jobs in entire sectors while creating new jobs in new sectors.[76] Generative AI is a potentially extinction-level event, but only for those who fail to deploy it strategically in their arsenal.[77]

Deloitte's research into content supply chains shows that leaders in content marketing who heavily utilize generative AI for creating content tend to be more successful.[78] The demand for content has surged by 54% on average in the past year, yet companies can only fulfill these demands about 55% of the time.[79] Generative AI is capable of quickly producing a wide array of content types, from social media updates to comprehensive

71. *Id.*

72. *Id.*

73. *Id.*

74. *Id.* ("The Abba avatars that sell out a London arena seven times a week and the celebrity-voiced chatbots recently launched by Meta are just a taste of how the biggest stars will be able to satisfy their fans.").

75. *Artificial intelligence is remixing journalism into a "soup" of language*, THE ECONOMIST (May 4, 2023), https://www.economist.com/business/2023/05/04/artificial-intelligence-is-remixing-journalism-into-a-soup-of-language [https://perma.cc/J899-433U] (reporting that Gina Chua Semafor executive editor declared, "We are not here to save journalists; we are here to save journalism.").

76. Tim W. Dornis, *Artificial Creativity: Emergent Works and the Void in Current Copyright Doctrine*, 22 YALE J. L. & TECH. 1, 40 (2020) ("structural changes in labor markets must be moderated by lawmakers, but they cannot be stopped.").

77. *See*, *e.g.*, Benjamin Mullin & Katie Robertson, *BuzzFeed News, Which Dragged Media Into the Digital Age, Shuts Down*, NY TIMES (Apr. 20, 203), https://www.nytimes.com/2023/04/20/business/buzzfeed-news-shut-down.html; s*ee also* Reuters, *Vice Media reportedly preparing to file for bankruptcy*, CNBC (May 2, 2023, 3:11 AM), https://www.cnbc.com/2023/05/02/vice-media-reportedly-preparing-to-file-for-bankruptcy.html [https://perma.cc/TDN4-WTVG].

78. Mike Brinker, *3 ways generative AI is changing the world of content*, DELOITTE DIGITAL (Oct. 8, 2023), https://www.deloittedigital.com/us/en/insights/perspective/adobe-max-genai.html [https://perma.cc/QLV4-A49Z].

79. *Id.*

reports and extensive videos.[80] This technology alleviates the burden of repetitive, time-intensive tasks from human staff, allowing them to concentrate on strategic decisions and innovative concepts where their skills are indispensable.[81]

Moreover, generative AI is poised to significantly alter the traditional dynamics of the content creation landscape, affecting advertisers, strategic allies, and fundamental business operations. Companies that incorporate generative AI reported saving their content marketing teams an average of 12.2 hours per full-time employee each week and were able to satisfy content needs 66% of the time, which is over 20 percentage points more than companies at a lower level of maturity that don't plan to use generative AI.[82] Firms that employ generative AI for content creation are more likely to deem their content marketing strategies highly effective, with a stark contrast of 30% compared to just 8% among less advanced companies.[83] Furthermore, 63% of businesses have experienced revenue growth as a result of integrating generative AI into their content marketing strategies.[84] The adoption of real-time, automated capabilities ensures content that aligns with customer preferences and habits, leading to smoother brand interaction, higher engagement rates, and increased customer loyalty.[85] Enhanced content personalization through generative AI also reduces the need for additional follow-up actions by employees, as customers receive more relevant content from the outset.

AI can assess the newsworthiness of research papers to help a busy newsroom with an initial sift, scan social media for news ripples, and use regional data to localize the message for readers through language translation, dialect intricacies, and news integration. Radar AI's five human journalists have filed over 400,000 partly automated stories since 2018.[86] Photorealistic images are available for purchase from major stock photography agencies alongside traditional photos. The value of human-generated work may increase thanks to AI-generated art. As the owner of a Hermès Birkin bag knows, handmade items command a premium next to mass market-produced ones.[87]

80. *Id.*
81. *Id.*
82. *Id.*
83. *Id.*
84. *Id.*
85. *Id.*
86. *Id.*
87. *Birkin Bag*, WIKIPEDIA, https://en.wikipedia.org/wiki/Birkin_bag [https://perma.cc/K88F-Y4GM] ("Birkin bags are handmade from leather and are named after the English-French actress and singer Jane Birkin. … In 2020, retail prices start at US$11,400 for a Birkin 25 bag. Costs can vary

Generative AI has democratized creativity in a way never before seen. Anyone with an Internet connection immediately had access to a host of text generators, text-to-image generators, and, soon, text-to-video generators. A robotic school of artwork is emerging. Midjourney and Dall-E generate images in seconds using simple written prompts. Sora, by OpenAI, "creates videos that look as if they were lifted from a Hollywood movie."[88] No artistic skill is required.

Rather than using digital art techniques such as ray tracing, in which a computer calculates how light will fall on an object from a particular angle, AI imagines how things will look based on the images in its training data, illuminated from many different angles. For now, even photorealistic AI images seem more akin to painting than naturalistic photography. Alongside photographers will be artists who use cameras and AI images as an "interstitial space between drawings, photographs and sculpture."[89] How should the law then understand the relationship between AI and creatives?

2. Tools and Co-Creators

In the 2024 Super Bowl commercial by Microsoft, we see various individuals surmounting obstacles, such as launching a business or achieving academic success, all powered by Copilot, Microsoft's AI assistant designed for everyday use.[90] This advertisement offers a glimpse into a future where AI could become as ubiquitous as smartphones, the internet, or electricity.[91] Despite the enthusiasm of many technology proponents who herald AI as the future, a significant portion of the general public remains wary.[92] A Pew Research Study recently found that 52% of

widely according to the type of leather, if exotic skins are used, and if precious metals and jewels are part of the bag. A bag made of exotic skin and diamond was sold at auction by Christie's in Hong Kong for a record price of US$380,000 in May 2017.").

88. *See* Cade Metz, *OpenAI Unveils A.I. That Instantly Generates Eye-Popping Videos*, N.Y. TIMES (Feb. 15, 2024), https://www.nytimes.com/2024/02/15/technology/openai-sora-videos.html.

89. Gideon Jacobs, *AI Is the Future of Photography. Does That Mean Photography Is Dead?*, N.Y. TIMES (Dec. 26, 2023), https://www.nytimes.com/2023/12/26/opinion/ai-future-photography.html.

90. *See* Microsoft, *Microsoft Game Day Commercial | Copilot: Your Everyday AI Companion*, YOUTUBE (Feb. 10, 2024), https://www.youtube.com/watch?v=SaCVSUbYpVc.

91. *See* Michelle Castillo, *The Way AI Will Be Sold to A Skeptical Public Is Starting To Become Clear*, CNBC (Feb. 13, 2024), https://www.cnbc.com/amp/2024/02/13/the-way-ai-will-be-sold-to-a-skeptical-public-is-becoming-clear.html [https://perma.cc/S7T5-928Q].

92. *See* Alec Tyson & Emma Kikuchi, *Growing Public Concern About the Role Of Artificial Intelligence In Daily Life*, PEW RES. CTR. (Aug. 28, 2023), https://www.pewresearch.org/short-reads/2023/08/28/growing-public-concern-about-the-role-of-artificial-intelligence-in-daily-life/ [https://perma.cc/4ZAT-Q349].

Americans express more concern than excitement over the advancement of AI, with only 10% feeling more excitement than apprehension about its potential. [93]

Concerns that AI might displace human workers overlook historical and economic evidence to the contrary. Similar to how calculators have become indispensable tools without replacing mathematical skills, AI enhances the capabilities of well-trained and judicious workers to tackle tasks that are otherwise tedious, prone to error, or simply impossible.[94] This technology significantly boosts the performance of both the most and least efficient workers. [95] While the entertainment industry has grown in size and profitability due to copyright laws, the proportion of earnings received by creators has diminished both in real terms and relative to the industry's overall profits. [96]

Innovation continuously introduces new tools, many of which are adopted through automation. Generative AI is no exception, offering efficiency, enhanced performance, and broader applications. For example, navigation technologies have simplified the navigational skills required by Uber drivers, effectively creating a new form of employment. If innovation were solely about automation, job scarcity would have been an issue by now. However, the challenge seems to be a shortage of workers rather than a shortage of jobs. Generative AI will indeed automate some tasks, render certain jobs obsolete, and transform others. The real question is about the quality and desirability of these evolving jobs.

Generative AI has the potential to amplify the impact and applicability of human expertise. It can integrate knowledge and rules with experiential learning to aid in complex decision-making, often reserved for specialists. [97] When utilized effectively, AI has the potential

93. *Id.*

94. Shakked Noy & Whitney Zhang, *Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence*, 381 SCIENCE Issue 6654, 187, available at: https://www.science.org/doi/10.1126/science.adh2586 (July 13, 2023) (finding significant improvements in the speed and quality of writing output among those assigned to the ChatGPT group).

95. *Id*. (finding the least productive workers using ChatGPT were about as effective as the median writer without ChatGPT). "ChatGPT enabled the most capable of writing faster and the less capable to write both faster and better—so the productivity gap between adequate and excellent writers shrank." *Id.*

96. Megan Morrone, *Copyright law will shape how we use generative AI*, AXIOS (Dec. 1, 2023) https://www.axios.com/2023/12/01/ai-copyright-law-biden-executive-order [https://perma.cc/K43W-7ZKU].

97. *See* David Autor, *AI Could Actually Help Rebuild the Middle Class*, NOEMA (Feb. 12, 2024), https://www.noemamag.com/how-ai-could-help-rebuild-the-middle-class/ [https://perma.cc/JM6F-KYEC] ("The unique opportunity that AI offers humanity is to turn back this tide — to extend the relevance, reach and value of human expertise for a larger set of workers.").

to rejuvenate the middle-skill, middle-class segment of the American workforce, which has been eroded by automation and globalization. The focus should not be on what AI will do to us but rather on how we can harness AI to serve our purposes.[98]

III. GOVERNING INFRINGEMENT

AI datasets comprise a broad range of sources, including data obtained by scraping content from publicly available websites and platforms.[99] In some cases, the dataset comprises copies of such works, but in others, it may comprise metadata or links to data.[100] The use of copyrighted materials by AI for training and the potential solutions through private ordering mechanisms like licensing raises interrelated challenges and opportunities. The law needs to clarify when generative AI inputs and outputs are infringing and related and when fair use applies. As it does, private ordering will occur in parallel with transitioning industrial production to a new normal.[101]

A. Infringement by Ingestion

The New York Times initiated legal action against prominent generative AI firms, accusing them of "illegally utilizing The Times's material to develop AI products that rival" the newspaper's offerings.[102] The complaint claims that these companies have been developing their models and tools using content from The Times without authorization.[103] After efforts to negotiate a settlement reached an impasse over several months, The New York Times proceeded to file a lawsuit.[104] The newspaper contends that the use of its content by these companies to create products that replicate the training inputs without compensation falls outside the scope of fair use and thus infringes on copyright laws.[105]

98. *Id.*

99. Samir Sampat, *Where Do Generative AI Models Source Their Data & Information*?, SMITH.AI (Sept. 20, 2023), https://smith.ai/blog/where-do-generative-ai-models-source-their-data-information [https://perma.cc/28RS-QMCY].

100. *See* Jonathan E. Barbee, *Scratching the Surface of IP Rights: Data Scraping and Artificial Intelligence*, NY STATE BAR ASS'N (Dec. 22, 2023), https://nysba.org/scratching-the-surface-of-ip-rights-data-scraping-and-artificial-intelligence/ [https://perma.cc/SRM5-B6L4].

101. *See infra*, Part II.B.

102. N.Y. Times v. Microsoft Corp., No. 1:23-cv-11195, filed, (S.D.N.Y. Dec. 27, 2023).

103. Complaint and Demand for Jury Trial, N.Y. Times v. Microsoft Corp., No 1:23-cv-11195, filed (S.D.N.Y. Dec 27, 2023).

104. *Id.*

105. *Id.*

For creatives, generative AI evokes dark reminders of Napster, a platform at the turn of the millennium for sharing mainly pirated songs.[106] Like Napster, developers seem to blithely ignore copyright, hoovering content without consent, credit, or compensation.[107] Moreover, the data is often undisclosed, making it challenging for creatives to identify potential use of their materials.[108] Creatives warn regulators against encouraging the Napster approach—enabling rapid market deployment first and plugging the gaps later.[109] Authors sue when they discover that their works have been used to train generative AI models without their permission. They point to accurate summaries of their work when their work is listed as part of a dataset that is alleged to have been used in training the model or the model's output is similar to or reproduces portions of their existing works.[110]

Developers say generative AI takes inspiration from works, much like a human art student.[111] The more access to training sets, the better the output will be.[112] Generative AI models are designed not to copy expressive elements but to make use of facts and patterns to compose new and substantially different works from the individual copyrighted works

106. *See* Michael Beausoleil, *AI Is Going to Be Napster 2.0 for the Music Industry*, MEDIUM (Apr. 23, 2023) https://beausoleil.medium.com/ai-is-going-to-be-napster-2-0-for-the-music-industry-8fd0b2ac6318.

107. *Id.*

108. David Luyt, *Undisclosed AI Systems and Data Protection*, MICHALSONS (Jan. 19, 2023) https://www.michalsons.com/blog/undisclosed-ai-systems-and-data-protection/63559 [https://perma.cc/PUT3-6YCW].

109. David De Cremer, Nicola Morini Bianzino, & Ben Falk, *How Generative AI Could Disrupt Creative Work*, HARV. BUS. REVIEW (Apr. 13, 2023), https://hbr.org/2023/04/how-generative-ai-could-disrupt-creative-work [https://perma.cc/5L85-9PSZ] ("intellectual-property laws have not caught up with the technological progress made in the field of AI research. It's quite possible that governments will spend decades fighting over how to balance incentives for technical innovation while retaining incentives for authentic human creation — a route that would be a terrific loss for human creativity.").

110. Cheryl Seah, *Generative AI*, LAW GAZETTE, https://lawgazette.com.sg/feature/generative-ai/ [https://perma.cc/F5Y2-X5RM] ("it cannot be the case that every output from ChatGPT is an infringing derivative of any author's work – for example, a paragraph describing Homer's Iliad could not possibly have been extracted from an American's autobiography.").

111. *See* Matthew Sag, Testimony before the U.S. Senate Committee on the Judiciary Subcommittee on Intellectual Property on 12 June 2023, at 4, available at: https://www.judiciary.senate.gov/imo/media/doc/2023-07-12_pm_-_testimony_-_sag.pdf; *see also* Andrew W. Torrance, Bill Tomlinson, *Training Is Everything: Artificial Intelligence, Copyright, and "Fair Training"*, 128 DICK. L. REV. 233, 252 (2023) ("A similar example might involve written work made possible through a training set of existing, and copyrighted, writing.").

112. Sara Brown, *Machine learning, Explained*, IDEAS MADE TO MATTER: ARTIFICIAL INTELLIGENCE (Apr. 21, 2021), https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained [https://perma.cc/HAN6-D9XE].

that are part of its training set.[113] They do not digest content in words but employ "subword tokens" to determine the probability that one subword token will follow another to form words, phrases, and sentences.[114] It is inherently derivative and generates content by selecting words or pixels based on extensive training data. It does not copy wholesale but develops parameters from training materials, making exact or even substantially similar reproduction unlikely. They warn that without access to a deluge of works, copyrighted or otherwise, generative AI may perish in its infancy.

Courts have required authors to show that the output is substantially similar to the copyrighted works.[115] Presenting evidence of infringing works produced by AI tools substantially similar to their copyrighted material is a major issue because outputs rarely match the material used in the training data closely enough to be substantially similar.[116] The court in *Andersen et al. v. Stability AI Ltd*. likewise found searches on "haveibeentrained.com," a tool that allows searching of the Laion-5B and Laion-400M image datasets, plausibly established copyrighted works at issue were included in those datasets.[117] However, it was unclear whether the model contained actual copies of the training images or algorithms that could reconstruct the images, as the AI products did not allow users to create new works by expressly referencing the author's works by name.[118] The models behind image generators are trained on enormous

113. *See* 17 U.S.C. § 102(a) ("Copyright protection subsists . . . in original works of authorship fixed in any tangible medium of expression."); *see also* James Vincent, *The Scary Truth About AI Copyright is Nobody Knows What Will Happen Next*, VERGE (Nov. 15, 2022) https://tinyurl.com/56xtcusr [https://perma.cc/XD49-DEWV] ("If the [text-to-image] model is training on millions of images and used to generate novel pictures, it's extremely unlikely that this constitutes copyright infringement. The training data has been transformed in the process, and the output does not threaten the market for the original art.").

114. *AI Tokens: The Building Blocks of Language Models*, POVIO, https://povio.com/blog/ai-tokens-the-building-blocks-of-language-models/ [https://perma.cc/Q3BL-ZXM6].

115. *See generally* Daryl Lim, *Substantial Similarity's Silent Death*, 48 PEPP. L. REV. 713 (2021); Daryl Lim, *Saving Substantial Similarity*, 73 FLA. L. REV. 591 (2021).

116. Benj Edwards, *Stable Diffusion "Memorizes" Some Images, Sparking Privacy Concerns*, ARS TECHNICA (Feb. 1, 2023, 1:37 PM), https://arstechnica.com/information-technology/2023/02/researchers-extract-training-images-from-stable-diffusion-but-its-difficult/ [https://perma.cc/F9SR-EC3L]; Jooyoung Lee et al., *Do Language Models Plagiarize*?, available at: https://arxiv.org/pdf/2203.07618.pdf [https://perma.cc/G8U6-2LJA].

117. Andersen v. Stability AI Ltd., No. 23-CV-00201-WHO, 2023 WL 7132064, at *4 (N.D. Cal. Oct. 30, 2023).

118. *Id.* at *6.

datasets: LAION-5B, the largest publicly available one, contains 5.85 billion tagged images.[119]

Courts have also rejected as implausible that all images generated using the model constitute infringing derivative works of every image used to train the model unless the images generated through the model are substantially similar to the copyrighted works.[120] For instance, the court in *Kadrey v. Meta Platforms Inc.* stated that "there is no way to understand the LLaMA models themselves as a recasting or adaptation of any of the plaintiffs' books."[121] To show outputs infringed derivative work rights, the outputs had to "incorporate in some form a portion of" the copyrighted books and be substantially similar.[122] Well, and good, but what about the input?

Section 107 of the Copyright Act provides that the fair use of a copyrighted work … is not an infringement of copyright."[123] Using copyrighted works is typically fair when the use serves a valuable social purpose and transforms the source material.[124] Judicial decisions have established that search engines, such as Google and Bing, which index vast quantities of copyrighted content on the internet, are shielded by the fair use doctrine. This protection is granted because the thumbnail images and excerpts of text they present in search results are considered "transformative."[125] Creating wholesale copies of a work is a preliminary

119. Christoph Schuhmann et al., *LAION-5B: An Open Large-Scale Dataset For Training Next Generation Image-Text Models*, OPENREVIEW.NET, https://openreview.net/forum?id=M3Y74vmsMcY [https://perma.cc/2556-7AZ4].

120. Owen Wolfe, Lauren Gregory Leipold & Grayson Moronta, *"The Plaintiffs Are Wrong": OpenAI Submits New Authority In Attempt To Knock Out Sarah Silverman's Claims* MONDAQ (Dec. 1 2023), https://www.mondaq.com/unitedstates/copyright/1397058/the-plaintiffs-are-wrong-openai-submits-new-authority-in-attempt-to-knock-out-sarah-silvermans-claims [https://perma.cc/Y5CE-976R].

121. Kadrey v. Meta Platforms, Inc., No. 23-CV-03417-VC, 2023 WL 8039640, at *1 (N.D. Cal. Nov. 20, 2023).

122. *Id.*

123. Sancton v. OpenAI Inc., et al., 1:2023-cv-10211, filed, (S.D.N.Y. Nov. 21, 2023).

124. Menand, *supra* note 62 ("[I]n Authors Guild v. Google, decided in 2015, courts upheld the legality of Google Books, even though it is a Web site that was created by scanning tens of millions of books without permission from the copyright holders. That case didn't even go to trial. Google won in summary judgment under the principle of fair use, and an appeals court held that Google Books' copying had a "highly convincing transformative purpose" and did not constitute copyright infringement. The outcome portends trouble for parties with copyright cases against companies that use A.I.").

125. *See e.g.,* Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2003); Perfect 10, Inc. v. Amazon.com, Inc., 508 F.3d 1146 (9th Cir. 2007).

step to developing a new, non-infringing product, even if the new product competes with the original. [126]

A permissive view focuses on the transitory and non-consumptive use that does not materially interfere with owners' copyrights and, therefore, should be fair use. [127] Copyright safeguards books from unauthorized reproduction and sale by others as substantially similar works. [128] Yet, within the vast expanse of training data for an algorithm, a single book becomes merely a collection of tokens akin to a drop in the ocean. [129]

On the creative side, in theory, creatives can use technical such as robot.txt to indicate their preferences known to AI systems. [130] Individuals who prefer to prevent their works from being used in AI training could implement a digital marker, akin to a website's indexing tag, signaling AI systems to exclude their content. [131] While the option to exclude works from AI training data exists and may become more common, it currently involves significant technical and practical challenges.

On the developer's side, the large quantity of work training AI needs makes obtaining authorization a significant burden, if not impracticable. [132] In the early stages of the internet, creative works were stored as distinct content pieces, which could be taken down if they were

126. Google LLC v. Oracle Am., Inc., 141 S. Ct. 1183 (2021) (holding that Google's use of a portion of the Oracle Java computer program in Googles Android operating system constituted fair use even though the defendant used the copyrighted content in a commercial venture after having tried and failed to negotiate a license that would have encompassed the content.).

127. Copyright Act of 1976, 17 U.S.C. § 107. Courts apply each of the four enumerated factors. The "purpose and character of the use" considers whether copying furthers a business or commercial purpose, the nature of a copied work, the amount copied, and market effect on the copyrighted work.

128. Manish Jinda, *Anti-Piracy Protection for Authors and Publishers*, BYTESCAREBLOGS, (Dec. 5, 2023), https://bytescare.com/blog/anti-piracy-protection-for-authors-and-publishers [https://perma.cc/JEA9-E4YF].

129. Menand, *supra* note 62 ("What the models are "learning" are not even sentences but "tokens," which are often pieces of words. When functioning properly, a model predicts, based on a statistical calculation, what token comes next. This has been mocked as simply an advanced form of autofill. But, when I write a sentence, I, too, am trying to guess the best next word. It just doesn't feel especially "auto." One big difference is that, since I fancy myself a writer, I am trying to avoid, wherever possible, the statistically most common solution.").

130. *Introduction to robots.txt*, GOOGLE SEARCH CENT., https://developers.google.com/search/docs/crawling-indexing/robots/intro [https://perma.cc/5Q3P-5AWB] (last visited May 10, 2024) ("A robots.txt file tells search engine crawlers which URLs the crawler can access on your site.").

131. *See generally Block Search Indexing with 'Noindex'*, GOOGLE SEARCH CENT., https://developers.google.com/search/docs/crawling-indexing/block-indexing [https://perma.cc/LN6H-WJWP] (last visited May 10, 2024) (describing how users can use a noindex meta tag "to prevent indexing content by search engines that support the noindex rule").

132. *See* INNOVATION, SCI. & ECON. DEV. CANADA, A CONSULTATION ON A MODERN COPYRIGHT FRAMEWORK FOR ARTIFICIAL INTELLIGENCE AND THE INTERNET OF THINGS 8 (2021).

found to violate copyright laws.[133] Contemporary AI technology operates differently. Creative works contribute to the training of AI models, where each piece influences minute adjustments distributed throughout the vast network of values, as seen with OpenAI's GPT-4, which involves changes across trillions of parameters.[134] It is not possible to "remove" the work that the model has already been trained on. Training data influences model weights, just as it would not be possible to identify and remove a piece of chicken blended with other parts into a chicken nugget.[135] To eliminate infringing content from generative AI systems such as ChatGPT or MidJourney, the sole method is to retrain the AI entirely from the beginning, excluding the contentious item(s).[136]

Major corporations, academic institutions, and individuals do not have access to similar levels for clearance. Fair use could reduce the economics of scale larger entities enjoy over smaller businesses in litigating and obtaining rights clearances. Making AI training available to organizations and individuals can "ensure diverse contributions to the development of future AI systems."[137] Diverse datasets, models, and apps, in turn, better target inequity in AI deployment.[138] As Andrew Torrence and Bill Tomlinson put it,

> [t]he use of copyrighted works as training data is crucial for the development of AI, as it allows the algorithms to learn, understand, and improve upon the information they are processing. When the AI algorithm uses this training set, they are transforming the data into new and unique forms of knowledge and not producing copies of the original

133. SECTION 512 OF TITLE 17: A REPORT OF THE REGISTER OF COPYRIGHTS (2020), https://www.copyright.gov/policy/section512/section-512-full-report.pdf [https://perma.cc/S9SR-REFC], (Hereinafter Copyright Office Report); "Traditional content industries faced what many came to view as an existential threat, from the convergence of newly dominant and near-lossless digital media formats with a world-wide, interconnected network that facilitated the distribution of digital files of any type."). *Id.* at 8.

134. William S. Morriss, *Will AI Destroy the DMCA Copyright Compromise?*, LEXOLOGY (Feb. 14, 2024), https://www.lexology.com/library/detail.aspx?g=a21e358e-0738-4a4d-9654-32c75115b96d&utm_source=Lexology+Daily+Newsfeed&utm_medium=H%E2%80%A6 [https://perma.cc/QEQ9-ZDZH].

135. *See generally* Jie Xu et al., *Machine Unlearning: Solutions and Challenge*, 14 J. OF LATEX CLASS FILES 8, available at https://arxiv.org/pdf/2308.07061.pdf [https://perma.cc/L24Q-2Y95]. *See also* Cheryl Seah, *Generative AI*, LAW GAZETTE, https://lawgazette.com.sg/feature/generative-ai/ [https://perma.cc/NQW2-DFZW] ("Research into "machine unlearning" to selectively remove data points without full retraining is still ongoing, and it is very much model-dependent.").

136. Stephen Pastis, *AI's Un-Learning Problem: Researchers Say It's Virtually Impossible to Make An A.I. Model 'Forget' The Things It Learns From Private User Data*, FORTUNE (Aug. 30, 2023, 12:43 PM), https://fortune.com/europe/2023/08/30/researchers-impossible-remove-private-user-data-delete-trained-ai-models/ [https://perma.cc/C3KV-TDHS].

137. *See* Torrance & Tomlinson, *supra* note 11111, at 254.

138. *See* Daryl Lim, *AI, Equity, and the IP Gap*, 75 S.M.U. L. REV. 815, 845–55 (2022).

> works. Because the AI algorithms are transforming the original work, this use should not be considered a violation of the creators' copyright. Instead, such uses by AI algorithms should be protected under a "fair training" exception.
>
> In this light, "fair training" becomes a necessary concept for the democratization and continued development of AI. The "fair training" exception will balance the rights of copyright owners with the AI's ability to learn and grow.[139]

A stricter view sees training as unauthorized misappropriation; usurping a licensing market for copyright owners has been usurped.[140] The training materially interferes with owners' rights and is neither transformative nor non-consumptive use.[141] Permitting training would lead to market harm, as AI-generated outputs compete with or replace the original works.[142] Favoring authors would establish an obligation for developers to obtain prior permission and paid licensing for using copyrighted material in AI models.

Considering the immense expense of retraining such models, which can amount to millions of dollars, a DMCA-style notice and takedown approach is impractical, especially with the likelihood of receiving numerous takedown requests from various artists, authors, publishers, and rights holders.[143] Conversely, in the absence of DMCA-style legal safeguards, it's conceivable that current AI products, many of which seem to be trained on datasets containing substantial volumes of unlicensed content from third parties, could be overwhelmed and possibly crippled by extensive legal actions. Amidst this dilemma of content ingestion, interim solutions through private arrangements are crucial until a more permanent legal framework is established.[144]

---

139. Torrance & Tomlinson, *supra* note 111, at 245 ("We argue that training AI algorithms should also be considered a transformative use and therefore, should be considered "fair training.").

140. *See* Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith, 598 U.S. 508, 533 (2023) (scrutinizing the specific use of the copyrighted work rather than focusing on the transformative value of the allegedly infringing work.).

141. *See, e.g.,* James Vincent, *Getty Images Sues AI Art Generator Stable Diffusion in the US for Copyright Infringement*, VERGE (Feb. 6, 2023, 11:56 AM), https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion [https://perma.cc/9D84-JA3H] (discussing the claims in the Getty Images lawsuit including copyright infringement and transformative use arguments).

142. Torrance & Tomlinson, *supra* note111, at 247.

143. Chuan Li, *OpenAI's GPT-3 Language Model: A Technical Overview*, LAMBDA LABS (June 3, 2020), https://lambdalabs.com/blog/demystifying-gpt-3 [https://perma.cc/4NYQ-R45N].

144. Bruce D. Sokler, et al, *(Un)fair Use? Copyrighted Works as AI Training Data — AI: The Washington Report*, LEXOLOGY https://www.lexology.com/library/detail.aspx?g=fc73a939-513d-4595-82da-

B. The Promise of Private Ordering and Its Limits

1. Blanket Licensing and Micro Licensing

The co-dependence between content owners and AI developers grows as users demand high-quality, accurate outputs.[145] Shutterstock is creating AI tools with a Contributor Fund to compensate artists fairly.[146] YouTube's Content ID system offers a model where copyrighted content is used in return for sharing ad revenue with the original creators.[147]

Adobe's Firefly tools are reported to have been developed exclusively using materials from the public domain and those for which they have proper licenses.[148] It stands out as the only prominent stock image platform that accepts submissions of AI-generated images, which, according to some metrics, are surpassing the performance of images created by humans.[149]

Nonetheless, the Firefly model has attracted criticism. A significant number of contributors to Adobe Stock, which boasts a library of 300 million images, illustrations, and other content used to train Firefly, have

---

c83b9c64b444&utm_source=Lexology+Daily+Newsfeed&utm_medium=HT%E2%80%A6 ("Aside from government agencies, Congress, and the courts, the ultimate solution to the issues posed by the use of copyrighted works as AI training data may come from the marketplace.").

145. Benjamin Mullin, *Inside the News Industry's Uneasy Negotiations With OpenAI*, N.Y. TIMES, https://www.nytimes.com/2023/12/29/business/media/media-openai-chatgpt.html (Dec. 29, 2023) ("AI models consume so much energy because of the vast amount of data that the model is trained on, the complexity of the model, and the volume of requests made to the AI by users.").

146. *AI-Generated Content on Shutterstock: Contributor FAQ*, SHUTTERSTOCK (Apr. 18, 2024), https://support.submit.shutterstock.com/s/article/Shutterstock-ai-and-Computer-Vision-Contributor-FAQ?language=en_US [https://perma.cc/L7Y8-EPRA].

147. *How Content ID Works*, YOUTUBE https://support.google.com/youtube/answer/2797370?hl=en [https://perma.cc/QC77-4DAB] (last visited Jan. 20, 2024).

148. Sharon Goldman, *Adobe Stock Creators Aren't Happy With Firefly, The Company's 'Commercially Safe' Gen AI Tool*, VENTUREBEAT (June 20, 2023, 3:00 AM) https://venturebeat.com/ai/adobe-stock-creators-arent-happy-with-firefly-the-companys-commercially-safe-gen-ai-tool/ [https://perma.cc/JTG4-UH3X] ("Bradford Newman, who leads global law firm Baker McKenzie's machine learning and AI practice in its Palo Alto office, said Adobe's "commercially-safe" execution and indemnification offer is one of the first and the cleanest that he has seen — because Firefly was trained on Adobe Stock imagery provided by creators, and which Adobe says it has the ability to use for this purpose according to its Stock Contributor license agreement.").

149. Matt Growcoot, *AI Images Are Outperforming Photos on Adobe Stock* (June 6, 2023) https://petapixel.com/2023/06/06/ai-images-are-outperforming-photos-on-adobe-stock/ [https://perma.cc/NGM3-YLWE] ("Stock Performer, an analytic website for stock photographers, compiled the data in a recent blog post where it found that the average revenue per image per month (RPI/m for short) for AI images is a whopping four and a half times bigger than photos or any other format. Stock Performer found that the RPI/m of any media file on Adobe Stock was averaging 3.75 cents. But AI-generated images are returning a much healthier RPI/m of 17 cents. Furthermore, the revenue per download (RPD) for AI images is $1.94 compared to any other file which is $1.22.").

voiced concerns that Adobe utilized their stock images for Firefly's training without clear communication or explicit permission.[150] Moreover, these contributors argue that the influx of AI-generated images into Adobe Stock is undermining the value of the platform, with Firefly's rising popularity diminishing the likelihood of traditional stock images being purchased.[151] Adobe's intention with integrating generative AI like Firefly is to empower creators to capitalize on their creative skills.[152]

The terms of service for Adobe are notably expansive, giving Adobe the ongoing right to utilize stock images for any individual purchases and modifications in Photoshop across any medium, including those not yet conceived.[153] Despite the legality of Adobe's actions falling within their contractual rights, creative contributors contend that a more ethical approach would have involved informing Adobe Stock artists about the use of their work in training Firefly AI and providing them with an option to opt-out from the outset.[154]

Creatives could pool their works to produce training sets available for a fee as a blanket license, similar to an ASCAP license for music.[155] Creatives receive payment based on the number of images they submitted and "the number of licenses that those images generated [in 12 months]", with the bonus weighted towards the licenses.[156] Calculating royalties based on downloads may be doable for images because images are tagged with text descriptions when they are used to train text-to-image generators.[157] It is less likely to be able to do this for text, given how books, articles, etc., are used to train LLMs to recognize patterns across words and are not tagged the way images are.

Martin Senftleben suggested a levy system focused on the outputs of AI to be uniformly imposed on all generative AI system providers in the

150. Goldman, *supra* note 148 (reporting legal experts view that Adobe gets a license for perpetuity, for whatever medium shall be invented, and images generated using a model is a derivative of the billions of images in the dataset would likely not infringe on an artist's copyright).

151. *Id*.

152. *Id*.

153. *Id*.

154. *Id*.

155. *See generally* ASCAP Licensing, ASCAP: Frequently Asked Questions, AM. SOC'Y OF COMPOSERS, AUTHORS, & PUBLISHERS, https://tinyurl.com/mr2rk46m [https://perma.cc/7CP2-BBDJ] (last visited May 10, 2024) (describing how ASCAP licensing works).

156. *See Firefly FAQ for Adobe Stock Contributors*, ADOBE (Oct. 18, 2023), https://helpx.adobe.com/sg/stock/contributor/help/firefly-faq-for-adobe-stock-contributors.html [https://perma.cc/T9QW-W5FE] (describing the bonus compensation to be based on "eligible online images and their downloads," and subsequent bonuses to be "based on new approved images and downloads").

157. Cheryl Seah, *Generative AI*, LAW GAZETTE (Nov. 2023), https://lawgazette.com.sg/feature/generative-ai/ [https://perma.cc/B9V5-Q77R].

EU whenever their systems generate content that could potentially replace human-created works.[158] Christoph Geiger warns that Senftleben's approach is distinctly European in its essence, especially because its rationale for redistributing income may not seamlessly align with every copyright tradition, particularly that of the United States.[159]

A blanket license, possibly with fair, reasonable, and nondiscriminatory rates, could be workable. However, it may bring with it its own sets of antitrust challenges.[160] The intersection of patent law and antitrust regulation offers insight here. Patent holders in standard-setting organizations often agree to license essential patents on FRAND terms.[161] While the Patent Act doesn't mandate this, market pressures in networked technologies can be significant.[162] For instance, if a patent holder refuses to commit their technology to an industry standard, the standard-setting organization may choose an alternative standard.[163] In *Qualcomm*, the Ninth Circuit's decision underscores the preference for private ordering over antitrust enforcement, even in scenarios with apparent anticompetitive behavior and consumer harm.[164]

A micro-licensing approach akin to Spotify's model, could be feasible if metadata enabling the tracking of rights ownership is in place.[165] In a collaborative effort, global news publisher Axel Springer has

158. Martin Senftleben, *Generative AI and Author Remuneration*, 54 INT'L REV. INTELL. PROP. COMP. L. 1535 (2023) ("All providers of generative AI systems are equally exposed to the levy payment obligation the moment they offer their products and services in the EU."). Senftleben refers to the theory of the "domain public payant" ("paying public domain"), according to which the exploitation of the public domain should at least partly serve the living generation of authors. *Id.*

159. Christophe Geiger, *To Pay or Not to Pay (for Training Generative AI), That is the Question*, JOTWELL (December 18, 2023), https://ip.jotwell.com/to-pay-or-not-to-pay-for-training-generative-ai-that-is-the-question/ [https://perma.cc/928Z-QBE5] (reviewing Senftleben: "From a European point of view, however, the proposal is certainly compatible with a tradition of remunerated exceptions, as there is an established practice and case law about the distribution rules in favor of creators of this kind of remuneration via collective management organizations.").

160. Daryl Lim, *Standard Essential Patents, Trolls, and the Smartphone Wars: Triangulating the End Game*, 119 PENN ST. L. REV. 1 (2014).

161. *See generally* Norman V. Siebrasse & Thomas F. Cotter, *The Value of the Standard*, 101 MINN. L. REV. 1159 (2017).

162. Herbert Hovenkamp, *FRAND and Antitrust*, 105 CORNELL L. REV. 1683 (2020).

163. *See* D. Scott Bosworth et al., *FRAND Commitments and Royalties for Standard Essential Patents*, *in* COMPLICATIONS AND QUANDARIES IN THE ICT SECTOR: STANDARD ESSENTIAL PATENTS AND COMPETITION ISSUES 19, 26 (Ashish Bharadwaj, Vishwas H. Devaiah & Indranath Gupta eds., 2018).

164. *See* Fed. Trade Comm'n v. Qualcomm Inc., 969 F.3d 974, 1003 (9th Cir. 2020) (declin[ing] to "ascribe antitrust liability in . . . dynamic and rapidly changing technology markets without clearer proof of anticompetitive effect").

165. James Vincent, *The Scary Truth About AI Copyright Is Nobody Knows What Will Happen Next*, THE VERGE, (Nov. 15, 2022, 10:00 AM), https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data [https://perma.cc/LQZ9-NLC2] (interviewing

entered into an agreement with OpenAI, which will enable ChatGPT to provide summaries of Axel Springer's content in response to user inquiries.[166] This arrangement will see ChatGPT offering summaries of pertinent news articles from Axel Springer's portfolio, such as Politico, Business Insider, Bild, and Welt, whenever users pose questions.[167] These summaries will acknowledge the source by citing the Axel Springer publication and include a link to the complete article, with the aim of boosting traffic and increasing subscription income for Axel Springer's outlets.[168] Moreover, OpenAI will compensate Axel Springer for utilizing its content, including archived materials, to enhance the training of the large language models that underpin ChatGPT. The Associated Press has also agreed to license its reports for use by ChatGPT, and there are additional licensing agreements that have been finalized or are currently being negotiated.[169]

2. Glaze and Nightshade

Artists whose unique styles are deliberately imitated face not just a decrease in commissions and fundamental earnings but also the proliferation of inferior synthetic replicas online, which can tarnish their brand and reputation. Additionally, for some artists, their distinctive style is deeply intertwined with their identity. Glaze allows creatives to tweak pixels in ways indiscernible to humans but which make a digitized artwork appear dramatically different from an algorithm.[170] A Glaze enhancement called Nightshade is an offensive data poisoning tool, a companion to defensive style Glaze.[171] The higher the number of

Matthe Butterick, a lawyer representing plaintiffs in lawsuits against companies for scraping data to train AI models.).

166. Shirin Ghaffary, *OpenAI In Talks With Dozens of Publishers to License Content* (Jan. 4, 2024, 4:50 PM) https://www.bloomberg.com/news/articles/2024-01-04/openai-in-talks-with-dozens-of-publishers-to-license-content?sref=P6Q0mxvj.

167. *Id*.

168. CHATGPT, *OpenAI partners with major news publisher Axel Springer*, (Dec. 2023), https://community.openai.com/t/openai-partners-with-major-news-publisher-axel-springer/557375 [https://perma.cc/KY88-HY8A].

169. Matt O'Brien, *ChatGPT-Maker Openai Signs Deal With AP To License News Stories*, ASSOCIATED PRESS (July 13, 2023, 11:41 AM), https://apnews.com/article/openai-chatgpt-associated-press-ap-f86f84c5bcc2f3b98074b38521f5f75a [https://perma.cc/FD6P-KAZB].

170. AFP, *Artists Use Tech Weapons To Thwart AI Copycats*, THE DAILY STAR (Dec. 24, 2023, 10:22 AM), https://www.thedailystar.net/news/world/news/artists-use-tech-weapons-against-ai-copycats-3501796 [https://perma.cc/9E3H-55L3].

171. Thomas Claburn, *How Artists Can Poison Their Pics with Deadly Nightshade To Deter AI Scrapers*, THE REGISTER, (Jan. 20, 2024, 10:05 AM), https://www.theregister.com/2024/01/20/nightshade_ai_images/#:~:text=Described%20in%20a%20research%20paper,gets%20ingested%20for%20model%20training [https://perma.cc/F7U4-JB26]

"poisoned" images in the training data, the greater the disruption.[172] An image of a dog would be interpreted as a cat. Unpredictable responses of this sort make text-to-image models significantly less useful.[173] Moreover, Nightshade damages the algorithm.[174] Nightshade does have limitations. Processed images may be different from the original, particularly if artwork uses "flat colors and smooth backgrounds."[175] It also would "require 8 million images 'poisoned' to have a significant impact on generating similar images for LAION models."[176]

3. Indemnities and Public Relations

The risk of infringement extends to both developers and users of AI tools, prompting companies like Adobe, Shutterstock, OpenAI, IBM, Google, Microsoft, and Anthropic to implement IP indemnity commitments. Anthropic's Commercial Terms of Service, for example, indemnifies its enterprise Claude API customers from copyright infringement claims made against them for "authorized use our services or their outputs."[177] These indemnities provide a layer of security,

("Nightshade is computed as a multi-objective optimization that minimizes visible changes to the original image.").

172. *University Of Chicago Researchers Develop Technique To Poison Generative AI Image Scraping Tool Seen As A Way To Protect Artists' Copyright*, CAMPUS TECHNOLOGY (Nov. 6, 2023), https://campustechnology.com/articles/2023/11/06/university-of-chicago-researchers-develop-technique-to-poison-generative-ai-image-scraping.aspx#:~:text=%22Surprisingly%2C%20we%20show%20that%20a,bleed%20through%22%20to%20similar%20prompts [https://perma.cc/ZK3P-956H] ("A moderate number of Nightshade attacks can destabilize general features in a text-to-image generative model, effectively disabling its ability to generate meaningful images").

173. *Nightshade: the 'data poisoning' tool boosting fightback against AI*, THE WEEK (Jan. 31, 2024), https://theweek.com/tech/nightshade-data-poisoning-tool-ai [https://perma.cc/BE29-A9LC] ("Nightshade can provide a powerful tool for content owners to protect their intellectual property against model trainers that disregard or ignore copyright notices, do-not-scrape/crawl directives, and opt-out lists.").

174. Ryan Morrison, *Artists Turn AI Against Itself With Tools Designed To Protect Their Work And Break The Algorithm — Here's How They Are 'Poisoning The Well'* TOM'S GUIDE (Dec. 18, 2023) https://www.tomsguide.com/news/artists-turn-ai-against-itself-with-tools-designed-to-protect-their-work-and-break-the-algorithm-heres-how-they-are-poisoning-the-well [https://perma.cc/D5XU-3R6B].

175. Thomas Claburn, *How Artists Can Poison Their Pics with Deadly Nightshade to Deter AI Scrapers*, THE REGISTER, (Jan 20, 2024) https://www.theregister.com/2024/01/20/nightshade_ai_images/#:~:text=Described%20in%20a%20research%20paper,gets%20ingested%20for%20model%20training. [https://perma.cc/B33P-5CK8].

176. *Id*.

177. Jonathan Mollod & Peter Cramer, *Anthropic Joins the Party, Offers Copyright Shield to Enterprise AI Customers*, PROSKAUER (Dec. 19, 2023), https://newmedialaw.proskauer.com/2023/

especially given the potential duration and uncertainty of test cases in courts. Microsoft's indemnity, for example, covers the use of tools based on its in-house models and those developed by OpenAI, but not third-party models. [178] They generally are not applicable to customers of free versions of GenAI tools or uses that violate use restrictions in the Terms or arise from "willful misconduct or violations of law."[179]

According to the FTC, however, such indemnities may breach copyright laws as unfair competition or deceptive practices. [180] Scholars have raised concerns about the FTC's critical view on indemnification.[181] They argue that non-expressive fair use of copyrighted works can counter market dominance. [182] Moreover, indemnifying end-users clarifies confusion over infringement claims, assigns responsibility for infringement risks to those in the best position to manage them, and encourages lawful generative AI use while deterring infringing activities, as indemnities would not cover unlawful uses. [183]

When compliance becomes uncertain, it can inadvertently discourage lawful behavior. [184] In an ambiguous regulatory landscape, businesses will likely struggle with legal compliance. [185] It's insufficient for these businesses to be in a position where they must await a judicial decision to determine the legality of practices like text and data mining amidst antitrust litigation. Complex and peculiar regulations lead to lengthy and expensive legal battles, necessitating extensive research and

12/19/anthropic-joins-the-party-offers-copyright-shield-to-enterprise-ai-customers/ [https://perma.cc/NA5M-FQEN].

178. *Id*.

179. *Expanded Legal Protections and Improvements to Our API*, ANTHROPIC (Dec. 19, 2023), https://www.anthropic.com/news/expanded-legal-protections-api-improvements [https://perma.cc/9WYH-RXUH].

180. U.S. Fed. Trade Commission, Comment Letter on Artificial Intelligence and Copyright Notice of Inquiry, Docket No. 2023-6 (Oct. 30, 2023), https://www.regulations.gov/comment/COLC-2023-0006-8630 [https://perma.cc/8NPJ-VKRM].

181. Pamela Samuelson, Christopher Jon Sprigman, & Matthew Sag, *The FTC's Misguided Comments on Copyright Office Generative AI Questions*, PATENTLY-O (Dec 11, 2023) https://patentlyo.com/patent/2023/12/misguided-copyright-generative.html [https://perma.cc/K6QR-LZBS].

182. *Id*.

183. *Id*.

184. *See* Rohit Chopra & Lina M. Khan, *The Case for "Unfair Methods of Competition" Rulemaking*, 87 U. CHI. L. REV. 357, 359 (2020) ("[T]he reliance on case-by-case adjudication yields a system of enforcement that generates ambiguity, unduly drains resources from enforcers, [privileges incumbents,] and deprives individuals and firms of any real opportunity to democratically participate in the process [of creating substantive antitrust rules].").

185. *See* Thomas A. Piraino, Jr., *A New Approach to the Antitrust Analysis of Mergers*, 83 B.U. L. REV. 785, 807 (2003) (stating that the rule of reason has "become so confusing that it preclude[s] antitrust practitioners from advising their clients as to the legality of particular conduct").

the involvement of costly experts.[186] This lack of clarity in the law hinders compliance and leads to additional complications. A minimally invasive regulatory approach seems most conducive to innovation and progress in this context.

Equally, if regulations focus only on immediate threats, they risk becoming outdated quickly due to technological change. In this vein, scholars have argued that the FTC prematurely concluded that using copyrighted materials, freely available online, for training machine learning models might constitute a copyright violation.[187] This assumption is overly hasty, considering that the legal nuances of AI training about copyright law are being actively debated in various federal copyright infringement lawsuits.[188] Furthermore, a more nuanced interpretation of fair use principles in the context of AI training could argue that, in many cases, such practices should be considered fair use.[189] The authority to decide what constitutes copyright infringement or fair use rightly falls within the judicial domain, not the FTC's jurisdiction.

Moreover, the FTC's comments fail to address the potential consequences on academic research and the development of technology in the private sector, particularly regarding the necessity of expressing consent to use text and data mining or machine learning to extract or derive noncopyrightable abstract information from copyrighted works.[190] From the FTC's perspective, imposing such a significant obstacle to computational analysis of texts, images, and sounds does not seem likely to foster competition. However, the AI market is diverse, involving large tech companies and smaller entities actively participating in various aspects of AI development. Should the FTC deem that consent is needed for AI training, it could create barriers that disproportionately affect smaller firms, hindering their competitive abilities. Compared to Large firms, by contrast, companies can more easily absorb the costs associated with obtaining permissions.[191]

186. *See* WILLIAM KOLASKY ET AL., ABA SECTION OF ANTITRUST LAW, CONTROLLING COSTS OF ANTITRUST ENFORCEMENT AND LITIGATION 1 (2012) (stating that the report was made as "a response to concerns" about both "the costs imposed on businesses by the American system of antitrust enforcement" and "the length of time required to resolve antitrust issues both in litigation and in enforcement proceedings")

187. Samuelson et al., *supra* note 181.

188. *Id.*

189. *See, e.g.*, Mark A. Lemley & Bryan Casey, *Fair Learning*, 99 TEX. L. REV. 743 (2021).

190. Samuelson et al., *supra* note 181.

191. *See generally*, Daryl Lim, *Determinants of Socially Responsible AI*, 8 GEO. L. & TECH. REV. (forthcoming 2024).

Addressing concerns about new technology that has immense promise and high potential risks necessitates a clear and humble approach. Copyright law is well-equipped to evaluate these risks and benefits, offering mechanisms for assigning liability and facilitating cost-benefit analyses. In comparison, the antitrust instruments at the FTC's disposal are less suited for this purpose. Implementing mandatory insurance could provide a more effective way to distribute these risks. The indemnification policies of major tech companies, which the agency has noted in the comments it filed with the Copyright Office,[192] are steps in this direction and should be viewed not as problematic but rather as successful examples of private regulation.

In short, the FTC should see generative AI as an asset and not a threat. Incorporating this technology into antitrust enforcement is a significant shift that could potentially enhance policy performance. Collaborative efforts between humans and AI are essential for defining competition-related parameters and refining outcomes. Additionally, predictive analytics can improve antitrust interventions by combining science and art. However, the challenge remains in balancing regulation with the dynamic nature of technological advancements and market needs.

OpenAI has posted positions as a "writing community specialist" and "creator community specialist," focused on acting as an "ambassador for OpenAI" among creatives.[193] As James Grimmelmann observes, generating goodwill and positive public relations could help reduce lawsuits.[194] Microsoft also tied up with American labor unions to prepare an "open dialogue" to understand how AI will shape the workforce in the future.[195] AFL-CIO President Liz Shuler hailed the collaboration as "a groundbreaking" and "historic" alliance.[196] The alliance will mainly work on information sharing with labor leaders about working AI technology trends, AI worker and student education, joint policy, and skills

192. FTC Comments, https://www.regulations.gov/comment/COLC-2023-0006-8630 [https://perma.cc/HPZ8-DEH6] ("Many large technology firms possess vast financial resources that enable them to indemnify the users of their generative AI tools or obtain exclusive licenses to copyrighted (or otherwise proprietary) training data, potentially further entrenching the market power of these dominant firms").

193. Michelle Cheng, *Is OpenAI Extending An Olive Branch To Creators And Writers?*, QUARTZ (Jan. 12, 2024) https://qz.com/is-openai-extending-an-olive-branch-to-creators-and-wri-1851156896 [https://perma.cc/359D-D9QP].

194. *Id.*

195. *Microsoft To Team Up with American Labor Unions To Discuss The Impact Of AI On Worker* LIVEMINT (Dec. 11, 2023), https://www.livemint.com/companies/news/microsoft-to-team-up-with-american-labour-unions-to-discuss-the-impact-of-ai-on-workers-artificial-intelligence-11702307607038.html [https://perma.cc/HV96-Z3RM]

196. *Id.*

development, and incorporate worker perspectives and expertise in developing AI.[197]

While AI will be a legal Wild West for some time, history shows that the market will find a way. Copyright law has proven versatile in adapting to new technologies. Content creators will figure out ways to get paid. It may impede the development of AI models for a while, but it will be a speed bump over the long term. The creative industry will take a period of private ordering to transition to a new normal. Beyond private disputes, countries have awakened to the immense potential of generative AI. They are scrambling to devise regulations that balance innovation with a direction driven by national interests and politics. The Article discusses this next.

IV. MODELS OF GOVERNANCE

The US's innovation-centric, market-driven approach, Europe's rights-focused, regulatory stance, and Asia's pro-innovation policies, particularly China's supportive stance towards AI-generated works' copyrightability, highlight different approaches leading economics have taken to AI governance. It also underscores the imperative for international collaboration and the establishment of global norms to effectively govern AI's borderless and rapidly evolving nature, advocating for dynamic legal frameworks that reconcile technological advancement with the protection of human creators' rights.

A. America's Market-Driven Approach

America is home to Meta, Amazon, Adobe, Google, Nvidia, and OpenAI, companies that are major AI developers. Its approach to AI governance is market-driven, guided by powerful undercurrents promoting innovation and protecting free speech. By and large, tech companies have leeway to police themselves on safety and security risks.[198]

President Biden released an Executive Order (EO) on AI in 2023 that seeks to advance a "coordinated, Federal Government-wide approach."[199] The guidance and recommendations do not have legal force but rather seek to include stakeholders' data and insights to inform legislation.

197. *Id.*

198. Remarks on Artificial Intelligence, 2023 DAILY COMP. PRES. DOC. 202300641 (July 21, 2023) (announcing that several AI makers, including Amazon, Anthropic, Google, Inflection, Meta, Microsoft, and OpenAI, had agreed to self-regulate their systems).

199. Exec. Order No. 14110, 88 Fed. Reg. 75191, 75191 (Oct. 30, 2023).

Regarding copyright issues, the EO directs the US Copyright Office to issue "recommendations to the President on potential executive actions relating to copyright and AI."[200] It gives the US Copyright Office an opportunity to update or possibly change its current position with respect to GAI works.

While individual members and committees have been introducing fragmented AI-related bills, the leadership in Congress has been aiming for an all-encompassing AI regulatory framework. Between September and December 2023, a bipartisan team of senators, headed by Majority Leader Chuck Schumer from New York, organized a sequence of "AI Insight Forums." These were private sessions involving diverse stakeholders to discuss significant governance challenges posed by AI.[201] These forums aim to cultivate the necessary expertise among Senate members to facilitate the creation of effective bipartisan AI legislation in a matter of "months."[202]

The few bills introduced in the Senate have mostly revolved around deepfakes and watermarking.[203] As seen from the earlier Parts, America's approach has been not being too prescriptive. Until a court rules or, less likely, a legislative amendment happens, it remains an open question how copyright law treats the ingestion of training data.

B. Europe's Right-Centered Approach

Creatives want their output licensed and developers to document and disclose their data sources. The EU's ambition to influence global dynamics, known as the "Brussels effect," adds a narrative of digital sovereignty to the regulatory landscape. A newly proposed exception for data mining allows copyright owners to choose to "opt out," enabling them to maintain complete control over their works for text and data mining purposes.[204] Several major collective rights management organizations have declared their intention to withdraw their entire

200. *See id.* at 75207.

201. Bruce D. Sokler, et al., *Leader Schumer Holds the Senate's Inaugural AI Insight Forum — AI: The Washington Report,* MINTZ (Sep. 21, 2023), https://www.mintz.com/insights-center/viewpoints/2191/2023-09-21-leader-schumer-holds-senates-inaugural-ai-insight-forum [https://perma.cc/M6YY-B4EU].

202. Emily Birnbaum & Alex Wayne, *Schumer Thinks Congress Can Regulate AI, Despite a Bad Record on Tech*, BLOOMBERG (Jul. 12, 2023), https://www.bloomberg.com/news/newsletters/2023-07-12/schumer-predicts-ai-legislation-in-months [https://perma.cc/8KRX-56W9].

203. Emilia David and Jess Weatherbed, *The EU AI Act Passed — Here's What Comes Next*, THE VERGE (Updated Mar. 13, 2024, 8:30 AM EDT), https://www.theverge.com/2023/12/14/24001919/eu-ai-act-foundation-models-regulation-data

204. 2019 O.J. (L 130) 92, 113–14.

catalog from being used in data mining for machine learning.[205] This move is expected to diminish the pool of training data for AI technologies notably. European analysts have pointed out that such an "opt-out" provision, or subjecting data mining to exclusive rights, could hinder the progress of generative AI technologies, potentially making the EU a less appealing environment for AI developers.[206]

In February 2024, European lawmakers doubled down on the side of creatives, ratifying a provisional agreement that paves the way for landmark legislation for AI that will likely have worldwide implications.[207] While previous drafts of the AI Act demanded that developers provide detailed summaries about the use of "copyrighted training data," the latest iteration of the AI Act expands this requirement, calling for transparency about all training data used, irrespective of copyright status.[208]

The AI Act enforces transparency about the use of training data as a means to enable copyright owners to assert their rights against their works being used for data mining.[209] The accompanying explanation to the December 2023 version of the AI Act clarifies that the EU regulations

205. Geiger, *supra* note 159.

206. *Id*. ("The need to obtain individual authorizations and manage remuneration payments for AI training constitutes an additional cost factor in the form of transaction costs and licensing fees. If the costs involved are too high, it will negatively impact the ability of the EU's AI sector to compete on the world market").

207. Shane Snider, *EU AI Act Takes Another Step Forward*, INFORMA (Feb. 15, 2024) https://www.informationweek.com/machine-learning-ai/eu-ai-act-takes-another-step-forward#close-modal [https://perma.cc/CF4T-LXVE].

208. Bruce D. Sokler, et al., *(Un)fair Use? Copyrighted Works as AI Training Data — AI: The Washington Report*, LEXOLOGY (Jan. 10, 2024), https://www.lexology.com/library/detail.aspx?g=fc73a939-513d-4595-82da-c83b9c64b444&utm_source=Lexology+Daily+Newsfeed&utm_medium=HT%E2%80%A6 [https://perma.cc/R3MR-2HN3].

209. S. Alex Yang & Angela Huyue Zhang *The Case for Regulating Generative AI Through Common Law* (Feb 15, 2024), https://www.project-syndicate.org/commentary/european-union-ai-act-could-impede-innovation-by-s-alex-yang-and-angela-huyue-zhang-2024-02?utm_source=twitter&utm_medium=organic-social&utm_campaign=page-posts-february24&utm_post-type=link&utm_format=16%3A9&utm_creative=link-image&utm_post-date=2024-02-15&s=03 [https://perma.cc/CD8D-9L7P] ("The hope is that this transparency requirement will facilitate compensation negotiations between content creators and AI firms."). US Representatives Don Beyer and Anna Eshoo, the Vice Chair and Co-Chair of the Congressional AI Caucus, introduced the AI Foundation Model Transparency Act seeking to implement something similar. *Id.* Under the proposed Act, developers of certain AI models may be required to disclose the sources of data used to train their models, including "information necessary to assist copyright owners or data license holders with enforcing their copyright or data license protections . . .." *See* Don Beyer, *Beyer, Eshoo Introduce Landmark AI Regulation Bill* (Dec. 22, 2023), https://beyer.house.gov/news/documentsingle.aspx?DocumentID=6052 [https://perma.cc/L2D7-QDWW].

permit "reproductions and extractions of works or other subject matters for text and data mining, under specific conditions."[210] According to these rules, the onus is on the copyright owner to actively choose to exclude their work from being used for data mining. Once a copyright owner has clearly indicated their intention to opt-out, providers of general-purpose AI models are required to seek permission from the copyright holders for text and data mining of such works.[211]

The Act is a bold omnibus risk-based initiative that may perish under the weight of its ambition. Creatives may be pleased to see this requirement, but questions may also remain about whether they find that this goes far enough to prevent infringement effectively. After all, as seen in Part II, AI cannot unlearn, and solutions would still need to be found as to the rightsholder being put in the same position as if the infringement never happened; that is how the 'status quo ante' for data already used could be even established.

The European Commission plans to establish an AI Office tasked with regulating general-purpose AI models, guided by advice from independent experts.[212] This office will devise methods to assess the performance of these models and keep track of potential risks.[213] However, there are concerns that even if entities like OpenAI were to submit their datasets for review, the AI Office might lack the necessary resources to examine these submissions thoroughly.[214]

Economist David S. Evans has highlighted that Europe is trailing in AI innovation, with both Europe and the UK significantly behind China in terms of patent grants and applications.[215] Between 2018 and the previous year, US investors injected $206.3 billion into the AI field, in contrast to China's $82.3 billion and the EU and UK's combined $37.5

210. OPEN FUTURE FOUNDATION, *Compromise proposal on general purpose AI models/general purpose AI systems* (last visited May 19, 2024), https://www.openfuture.eu/wp-content/uploads/2023/12/231206GPAI_Compromise_proposalv4.pdf [https://perma.cc/RM7N-T4R5].

211. *Id*.

212. Elizabeth Gibney, *What the EU's Tough AI Law Means For Research And ChatGPT*, NATURE (Feb. 16, 2024), https://www.nature.com/articles/d41586-024-00497-8 [https://perma.cc/PF9H-S88B].

213. *Id*.

214. *Id*. ("The demand to be transparent is very important," they say. "But there was little thought spent on how these procedures have to be executed.").

215. *Europe Could Reach Tentative Agreement on AI Act This Week*, PYMNTS (Dec. 6, 2023), https://www.pymnts.com/news/artificial-intelligence/2023/europe-could-reach-tentative-agreement-ai-act-this-week/ [https://perma.cc/FX94-GCP3] ("The EU and EU+U.K. shares are so tiny relative to the U.S. and China, however, that they point to far less innovative effort going into AI.").

billion.[216] The dominance of US and Chinese companies in the commercial AI sector is notable, with many choosing the UK for their European headquarters.[217] The global AI market was valued at just over $450 billion in 2022 and is expected to soar to more than $1,800 billion by 2030. It represents a substantial market opportunity that the EU risks missing out on.[218]

C. Asian Pro-Innovation Approaches

As seen in Part I, China has shown it highly supports recognizing copyrightability for AI-generated works. It enacted the world's earliest and most detailed regulations that cover all generative AI content services, requiring developers to train the models in accordance with the law, including not infringing the intellectual property rights of third parties[219] China's headwind is that it lags in foundational model building, facing obstacles related to data accessibility, hardware constraints, and talent migration to the US.[220] Moreover, obstacles like the chilling effect of its recent move to enforce 'common prosperity' remain, and US export controls will continue to contribute to China's tech industry limitations.[221]

---

216. *Id.*

217. *Id.*

218. Nedim Malovic, *Transparency Requirements In The EU Provisional Agreement On The Artificial Intelligence Act Influence The Copyright And Wider Tech Debate*, IPKAt (Dec. 15, 2023), https://ipkitten.blogspot.com/2023/12/guest-post-transparency-requirements-in.html [https://perma.cc/K7E7-3DSV]; *see also* S. Alex Yang & Angela Huyue Zhang, *The Case for Regulating Generative AI Through Common Law* PROJECT SYNDICATE (Feb. 15, 2024), https://www.project-syndicate.org/commentary/european-union-ai-act-could-impede-innovation-by-s-alex-yang-and-angela-huyue-zhang-2024-02?utm_source=twitter&utm_medium=organic-social&utm_campaign=page-posts-february24&utm_post-type=link&utm_format=16%3A9&utm_creative=link-image&utm_post-date=2024-02-15&s=03 [https://perma.cc/6AGA-CFQK] The EU's AI Act, whose broad mandate applies to all firms regardless of their specific industry sectors, combined with the pace of AI development and the competitive structure of the market, increases the likelihood of serious unintended consequences. *Id.* Consequently, the common-law system, which is based on case-by-case adjudication, may turn out to be a more appropriate institutional framework for regulating AI. *Id.*

219. Interim Measures for the Management of Generative Artificial Intelligence Services (promulgated by the Office of the Central Cybersecurity and Information Commission, July 10, 2023, rev'd May 23, 2023, effective Aug. 15, 2023), art. 7 No. 15 (EN) (ChinaLawTranslate) https://www.chinalawtranslate.com/en/comparison-chart-of-current-vs-draft-rules-for-generative-ai/.

220. Daniel Castro, Michael McLaughlin and Eline Chivot, *Who Is Winning the AI Race: China, the EU or the United States?*, CENTER FOR DATA INNOVATION (Aug. 19, 2019) https://datainnovation.org/2019/08/who-is-winning-the-ai-race-china-the-eu-or-the-united-states/ [https://perma.cc/65FA-SVVG].

221. Ryan Hass, *Assessing China's "Common Prosperity" Campaign*, BROOKINGS, (Sep. 9, 2021) https://www.brookings.edu/articles/assessing-chinas-common-prosperity-campaign/ [https://perma.cc/W9MR-9GW4].

Investors have been spooked by the speed and force of implementation, which has wiped billions of dollars off the market.

Japan and Singapore have legislative exceptions that allow data mining. While Japan has a data mining exception, the Agency for Cultural Affairs released a draft report on AI and copyright, which supports an interpretation of the copyright laws under which certain unauthorized uses of protected content to train AI models would constitute a copyright violation.[222] Singapore provides a copyright exception for making a copy of the copyrighted work for computational data analysis and for preparing the work for such analysis.[223] The exception permits commercial use of the copyrighted work and cannot be overridden by a contract as long as the copy is used in a permitted manner, the user supplies that copy to others only for 'verifying the results of the computational data analysis' and 'collaborative research or study relating to the purpose of [such] analysis, the user needs to have lawful access to the copyrighted work that is to be used.[224]

Peter Yu observed that the Singapore provision,

> provides clarity by offering carefully drafted definitions and helpful illustrations. It also provides detailed language, thereby facilitating a more politically palatable compromise between copyright holders and their supportive industries on the one hand and users and AI developers on the other…. When all of the features of the new Singaporean copyright exception for computational data analysis are taken together, the exception may provide a more attractive model for transplant or emulation, especially among countries that remain concerned about the pushbacks from local copyright industries, foreign investors, and their powerful and supportive governments.[225]

---

222. Article 30-4 of the Japanese Copyright Act, Feb. 2020, https://www.cric.or.jp/english/clj/cl2.html ("It is permissible to exploit a work, in any way and to the extent considered necessary…if it is done for use in data analysis. . .."); Bruce D. Sokler, et al., *(Un)fair Use? Copyrighted Works as AI Training Data — AI: The Washington Report*, LEXOLOGY https://www.lexology.com/library/detail.aspx?g=fc73a939-513d-4595-82da-c83b9c64b444&utm_source=Lexology+Daily+Newsfeed&utm_medium=HT%E2%80%A6 [https://perma.cc/R3MR-2HN3] ("Some have interpreted this provision to mean that AI developers can use copyrighted works to train AI models without requesting the permission of copyright holders."); Hiroyuki Omoto & Kaoru Yamada, *Japan Panel Pushes to Shield Copyrighted Work From AI Training*, NIKKEI ASIA (Dec. 21, 2023), https://asia.nikkei.com/Business/Technology/Japan-panel-pushes-to-shield-copyrighted-work-from-AI-training [https://perma.cc/KS5S-7R4A].

223. Section 244 Singapore Copyright Act. (effective Dec. 31, 2021), https://wipolex-res.wipo.int/edocs/lexdocs/laws/en/sg/sg175en.pdf [https://perma.cc/G96V-PMK6].

224. *Id.*

225. Peter K. Yu, *The Future Path of Artificial Intelligence and Copyright Law in The Asian Pacific*, 96 COMP. & LAW (2024) (forthcoming)

Moreover, "the law in Singapore will likely be more supportive of such development due to the existence of both the fair use provision and an explicit copyright exception for computational data analysis. The United States, by contrast, has only the fair use provision."[226] In reflecting on the approach toward favoring developers, Peter Yu explained that:

> The push for stronger protection of AI-generated works in Asia is unsurprising for several reasons. To begin with, policymakers in many Asian countries believe that AI technology will drive their countries' economic and technological developments and global competitiveness. To the extent that these policymakers believe that stronger protection for AI-generated works will help provide the incentives or attract investments needed to fuel innovation in the AI sector, they will actively push for new laws and policies that offer such protection. In China, for example, AI has featured prominently in the country's strategic plans for economic, social, and science and technology development. Particularly ambitious is the country's goal of "becoming the world's major artificial intelligence innovation center" by 2030.
>
> In addition, the past decades have seen external pressure repeatedly put on Asian countries to strengthen the protection and enforcement of intellectual property rights. The demands for intellectual property reforms from foreign investors and their supportive governments have also created the misimpression that the key goal of intellectual property rights is to protect investments, as opposed to generating incentives for creation. To some extent, the origins of modern intellectual property laws in Asia, regardless of their historical roots, are somewhat different from those of the laws in Europe and the United States. Now that Asian countries have made the difficult transition to an intellectual property regime that offers higher levels of protection and enforcement and that, to a large extent, mirrors Western intellectual property regimes, it may be difficult and too late to convince Asian policymakers that AI-generated works should remain unprotected and go into the public domain.[227]

Effective governance of generative AI necessitates collaboration, trust-building, and the creation of global standards among nations. Yet, the challenge lies in the differing value systems and geopolitical tensions that hinder the achievement of a unified approach. Regulatory bodies around the world are adopting diverse strategies to oversee AI, leading to a disjointed set of rules for a technology that inherently knows no borders. To chart a path forward, it might be beneficial for policymakers to draw

226. *Id.*
227. *Id.*

lessons from sectors subject to stringent regulation, like pharmaceuticals and nuclear energy. While not directly comparable to AI, the rigorous quality controls and procedures that ensure safety in these critical domains could provide valuable guidance.[228]

V. CONCLUSION

As we stand on the precipice of a new era in creativity and innovation, this Article offers a profound exploration into the complex interplay between generative artificial intelligence technologies and the established norms of copyright law. The document underscores the urgency of reimagining governance frameworks to adeptly navigate the challenges and opportunities presented by AI's burgeoning role in creative industries.

The conclusion drawn from this comprehensive analysis is clear: the legal and regulatory landscapes must evolve in tandem with technological advancements to ensure a harmonious balance between fostering innovation and protecting the intellectual property rights of human creators. The stark contrasts in legal approaches to AI-generated content, from the conservative perspectives of the US to the more liberal views of China, highlight the global disparity in readiness and response to these technological shifts.

To navigate this uncharted territory, this Article advocates for a collaborative approach to policymaking, one that engages a diverse array of stakeholders, from legal experts to technologists and creators. The goal is to forge adaptable, forward-looking legal frameworks that can accommodate the nuances of AI-driven creativity while safeguarding the bedrock principles of copyright law.

This Article calls for a concerted effort to craft governance models that are as innovative and dynamic as the technologies they seek to regulate. By doing so, we can ensure that the digital age remains a fertile ground for creativity, where innovation thrives, and the rights of all creators are upheld with equity and justice. This journey towards a new legal paradigm is not only necessary but imperative as we strive to harness

228. *Id*. ("In the future, many creative works will feature a mix of human and machine contributions rather than contributions from humans or machines alone. It will, therefore, be useful to start asking whether copyright norms at both the domestic and international levels need to be adjusted when the human creative input has been significantly reduced – say, from eighty or ninety percent to ten or twenty percent. In a hybrid world where human creators work alongside AI systems, it will be important for legislatures and courts to devote greater attention to studying the interface between humans and machines in the creative process.").

the full potential of generative AI within a framework that respects and protects the timeless value of human creativity.